

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 26, 2003, 00:55:20 ; Search time 3253 Seconds
(without alignments)
15871.004 Million cell updates/sec

Title: US-09-854-300-7

Perfect score: 1774

Sequence: 1 agctgagctccaccgcggt.....ataggcaagttctctctcag 1774

Scoring table: IDENTITY NUC %

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.man.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1662.8	93.7	2796	6	AX337968 Sequence
2	1659.6	93.6	2830	9	AX074264 Homo sapi
3	1280.8	72.2	1297	9	AF394689 Homo sapi
4	1205.2	67.9	1928	10	BC010477 Mus muscu
5	1203.6	67.8	1936	10	AB041548 Mus muscu
6	1106.8	62.4	1287	10	AF426411 Mus muscu
7	714.4	40.3	1871	9	AK027169 Homo sapi
8	674.4	38.0	169812	2	AL391315 Human DNA
9	674.4	38.0	169863	2	AL591467 Homo sapi
10	392.2	22.1	404	6	AX071980 Sequence
11	372.6	21.0	646	6	AX341568 Sequence
12	310	17.5	435	6	AX396961 Sequence
13	280.4	15.8	195263	2	AC115936 Mus muscu
14	278.4	15.7	167021	2	AC111461 Rattus no
15	268.8	15.2	1391	9	AB070023 Macaca fa
16	268.8	15.2	1395	9	AB070058 Macaca fa
17	266.8	15.0	1429	9	AK098524 Homo sapi
18	266.8	15.0	1430	9	BC022038 Homo sapi
19	266.8	15.0	1446	9	AF447589 Homo sapi
20	266.8	15.0	120984	9	AC006463 Homo sapi
21	256.4	14.5	1328	9	BC029264 Homo sapi
22	254.8	14.4	1355	6	AX061659 Sequence
23	253.2	14.3	1261	6	AX098654 Homo sapi
24	220.2	12.4	116739	2	AC114183 Rattus no
25	220.2	12.4	140346	2	AC109686 Rattus no
26	220	12.4	119995	10	AL731648 Mouse DNA
27	217.4	12.3	116739	2	AC114183 Rattus no
28	209	11.8	1200	6	AX083416 Sequence
29	209	11.8	1527	6	AX083426 Sequence
30	209	11.8	2001	9	BC019355 Homo sapi
31	207.4	11.7	2013	9	BC032328 Homo sapi
32	156.8	8.8	839	9	AK058116 Homo sapi
33	148.8	8.4	1452	10	AF171875 Mus muscu
34	148.8	8.4	1506	10	BC018199 Mus muscu
35	147.6	8.3	1153	9	AF155650 Human mRNA
36	147.6	8.3	1253	6	AR030960 Sequence
37	147.6	8.3	1257	6	AX083320 Sequence
38	147.6	8.3	1433	6	AX083330 Sequence
39	147.6	8.3	1445	9	BC017100 Homo sapi
40	140	7.9	118766	2	AC128696 Rattus no
41	108.4	6.1	5111	9	AB033040 Homo sapi
42	105.6	6.0	382	6	AX261677 Sequence
43	87.2	4.9	2369	3	M97204 Fruitfly go
44	84.8	4.8	2586	3	AY069169 Drosophil
45	79	4.5	2448	4	OCU132110 Oryctolag

ALIGNMENTS

RESULT 1
AX337968
LOCUS AX337968 2796 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 9 from Patent WO0194391.
ACCESSION AX337968
VERSION AX337968.1 GI:18128679
KEYWORDS human.
SOURCE Homo sapiens

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1

AUTHORS

Xue, H., He, A., Nguyen, D.B., Yao, M.G., Bandman, O., Burford, N.,
Tang, Y.T., Xu, Y., Hafalia, A., Azimzai, Y. and Wallia, N.K.

INTRACELLULAR SIGNALING PROTEINS

JOURNAL Patent: WO 0194391-A 9 13-DEC-2001;

Incyte Genomics, Inc. (US)

Location/Qualifiers

source

1. .2796

/organism="Homo sapiens"

/db xref="taxon.9606"

/note="Incyte ID No: 2791668B1"

BASE COUNT 821 a 548 c 620 g 807 t

ORIGIN

Query Match 93.7%; Score 1662.8; DB 6; Length 2796;

Best Local Similarity 99.4%; Pred. No. 0;

Matches 1690; Conservative 0; Mismatches 8; Indels 3; Gaps 2;

```
Qy 74 CCGAGGAGCTGCATCTGGGCAACCTGTGTGCTGACGCTACGCTCCCTGGCTCCGAGC 133
Db 28 CCGAGGAGCTGCATCTGGGCAACCTGTGTGCTGACGCTACGCTCCCTGGCTCCGAGC 87

Qy 134 TAGCTCGCAGCTCCCAAGCTCTCACTCCATTCCTTCCCACTGCGCGCACCTGCTCAAG 193
Db 88 TAGCTCGCAGCTCCCAAGCTCTCACTCCATTCCTTCCCACTGCGCGCACCTGCTCAAG 147

Qy 194 ACCAGGCTCTCCCAAGCTCTAGGAGGCGGTGCGGAGGCGCTAGGAGGCGCTAGGAGG 253
Db 148 ACCAGGCTCTCCCAAGCTCTAGGAGGCGGTGCGGAGGCGCTAGGAGGCGCTAGGAGG 207

Qy 254 GCGCGCGCATGGGCGCGCGCTGCGGCGCGGCTCTCTGCGCGCGTGTGCGGCTTTT 313
Db 208 GCGCGCGCATGGGCGCGCGCTGCGGCGCGGCTCTCTGCGCGCGTGTGCGGCTTTT 267

Qy 314 TCCAGATTGCTGGCATGTGCTCTCTGCTGCGCTGAGTCCGCGAGGCGCGGTTCCCGG 373
Db 268 TCCAGATTGCTGGCATGTGCTCTCTGCTGCGCTGAGTCCGCGAGGCGCGGTTCCCGG 327

Qy 374 GGGCTGAAGCAGTGTGGACCGCTTACCTCAACGTGTCTGCGGGGTTCCGCAACGGA 433
Db 328 GGGCTGAAGCAGTGTGGACCGCTTACCTCAACGTGTCTGCGGGGTTCCGCAACGGA 387

Qy 434 GTGAACCTGAGTGTGGAGCTGAGCGAGGAGGCGGTGACGCGCAGGACTGCGCGCTG 493
Db 388 GTGAACCTGAGTGTGGAGCTGAGCGAGGAGGCGGTGACGCGCAGGACTGCGCGCTG 447

Qy 494 GAGCTGTGGTGTGGGCTCTGCTACCGCGCGAGCGCGCGCGGCGCTTAAGCGCTGTAC 553
Db 448 GAGCTGTGGTGTGGGCTCTGCTACCGCGCGAGCGCGCGCGGCGCTTAAGCGCTGTAC 507

Qy 554 CCGCACAGAAATTTACCGTGCCTACCGTTTGGGGAAGCACCGGTGCAAGTCTTTGGTTG 613
Db 508 CCGCACAGAAATTTACCGTGCCTACCGTTTGGGGAAGCACCGGTGCAAGTCTTTGGTTG 567

Qy 614 GCGCTCATCCACGCGCGGCGGCTGACCTTCCGCGAGGAGTCCATCTGCGCTTATGAG 673
Db 568 GCGCTCATCCACGCGCGGCGGCTGACCTTCCGCGAGGAGTCCATCTGCGCTTATGAG 627

Qy 674 AGATGGGCGTCTGGAGCGCTCATCTTTAACTTCCCGCGGACCGCAATGAGGTCAATCCC 733
Db 628 AGATGGGCGTCTGGAGCGCTCATCTTTAACTTCCCGCGGACCGCAATGAGGTCAATCCC 687

Qy 734 ATGTCTACCGCGGTGCGAGTACATTTGTTGAATCATGATCGGCAATCTGAAAGGACA 793
Db 688 ATGTCTACCGCGGTGCGAGTACATTTGTTGAATCATGATCGGCAATCTGAAAGGACA 747

Qy 794 AAAATTTCTGCAATCTATTCAAGAGGCGATACAAGTGACAAATGGTCATAGAGTAGGAAA 853
Db 748 AAAATTTCTGCAATCTATTCAAGAGGCGATACAAGTGACAAATGGTCATAGAGTAGGAAA 807

Qy 854 AAACATGCGCTTGGTGAATCACTATTCAATTTTTTTTCTGTTCTGTTCTTTTATT 913
Db 808 AAACATGCGCTTGGTGAATCACTATTCAATTTTTTTTCTGTTCTGTTCTTTTATT 867

Qy 914 ATTACGGCGGCACTCTGGCTATTTTATCTTTATCTGCTCGAAGGCTACGAATGCA 973
Db 868 ATTACGGCGGCACTCTGGCTATTTTATCTTTATCTTTATCTGCTCGAAGGCTACGAATGCA 927
```

```
Qy 974 AGAGCTCAAGCAGGAGCAGAGGCAATTAAAGGCAGATGCTAAAAAGCTATTGGAGG 1033
Db 928 AGAGCTCAAGCAGGAGCAGAGGCAATTAAAGGCAGATGCTAAAAAGCTATTGGAGG 987

Qy 1034 CTTCAACTACGCACACTGAAACAGGAGACAGGAAATGGCCCTGATGGAGATAGTTGT 1093
Db 988 CTTCAACTACGCACACTGAAACAGGAGACAGGAAATGGCCCTGATGGAGATAGTTGT 1047

Qy 1094 GCTGTGTCATTGAAATTTGATATAAACCAATGATTTGGTACGCATCTTAAACGTGCAACCAT 1153
Db 1048 GCTGTGTCATTGAAATTTGATATAAACCAATGATTTGGTACGCATCTTAAACGTGCAACCAT 1107

Qy 1154 ATTTTCCATAAGACATGTTGACCCATGGCTTTTAAACACACAGACATTGGCCCATGTGC 1213
Db 1108 ATTTTCCATAAGACATGTTGACCCATGGCTTTTAAACACACAGACATTGGCCCATGTGC 1167

Qy 1214 AAATGTGACATCTCAAGCTTTGGGAATTTGGGATGTTGAAGATGGATCAGTGCT 1273
Db 1168 AAATGTGACATCTCAAGCTTTGGGAATTTGGGATGTTGAAGATGGATCAGTGCT 1227

Qy 1274 TTACAAGTCCCTGTATCCAATGAAATATCTAATAGTGCCTCTCCCATGAAGAGGATAAT 1333
Db 1228 TTACAAGTCCCTGTATCCAATGAAATATCTAATAGTGCCTCTCCCATGAAGAGGATAAT 1287

Qy 1334 CCGCAGGAGACCGCATCATCTGGATATCTTCAGTACAGGGAACAGATGAACCGCTCTG 1393
Db 1288 CCGCAGGAGACCGCATCATCTGGATATCTTCAGTACAGGGAACAGATGAACCGCTCTG 1347

Qy 1394 GAGGAACACGTGACGCTCAACAAATGAAAGTCTACAGCTGGTAAACCATGAAGCAAAATCT 1453
Db 1348 GAGGAACACGTGACGCTCAACAAATGAAAGTCTACAGCTGGTAAACCATGAAGCAAAATCT 1407

Qy 1454 GTGCGAGTGGATGTTTCTCATGTNGACAAACCCACTTTTGAAGAGACGAAACTCT 1513
Db 1408 GTGCGAGTGGATGTTTCTCATGTNGACAAACCCACTTTTGAAGAGACGAAACTCT 1467

Qy 1514 AATCAAGAGACTGCTGTTGAGAAATTAATCTTAAATCTGTGTAATAGAAACTTGA 1573
Db 1468 AATCAAGAGACTGCTGTTGAGAAATTAATCTTAAATCTGTGTAATAGAAACTTGA 1527

Qy 1574 ACCATTAGTAATAACAGAACTGCCAATCAGGCGCTAGTTTCTATTAAATAATGGATAAA 1633
Db 1528 ACCATTAGTAATAACAGAACTGCCAATCAGGCGCTAGTTTCTATTAAATAATGGATAAA 1587

Qy 1634 TTTAATAAATAAGAGTGATGATCTGAAAGTGTCTCAGATGACTAATATTATGATAGTAA 1693
Db 1588 TTTAATAAATAAGAGTGATGATCTGAAAGTGTCTCAGATGACTAATATTATGATAGTAA 1647

Qy 1694 AATGCTTAAATAATTTAACTGTTAACTTTTCCACCAACTCATTATAATATTTT 1753
Db 1648 A--TGCTTAAATAATTTAACTGTTAACTTTTCCCA--CAAACTCAITATAATATTTT 1704

Qy 1754 CATAGGCAAGTTTCTCTCTCAG 1774
Db 1705 CATAGGCAAGTTTCTCTCTCAG 1725
```

RESULT 2
AK074264

LOCUS

DEFINITION Homo sapiens cDNA FLJ23684 fis, clone HEP09821.
ACCESSION AK074264

VERSION AK074264.1 GI:18676818

KEYWORDS oligo capping; fis (full insert sequence).

SOURCE Homo sapiens hepatoma cell line:HepG2 cDNA to mRNA, clone_lib:HEP

clone:HEP09821.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1

AUTHORS Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,

Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,

Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.
 NEDO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 2830)
 Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T., Isogai, T.,
 Shibahara, T., Tanaka, T. and Nakamura, Y.
 Direct Submission
 Submitted (14-PEB-2002) Sumio Sugano, Institute of Medical Science,
 University of Tokyo, Laboratory of Genome Structure, Human Genome
 Center, Shirokane-dai, 4-6-1, Minato-Ku, Tokyo 108-8639, Japan
 (E-mail: cdnalm@ims.u-tokyo.ac.jp, Tel: 81-3-5449-5286,
 Fax: 81-3-5449-5416)
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing;
 Research Association for Biotechnology; cDNA library construction;
 5'- & 3'- end one pass sequencing; Department of Virology and Human
 Genome Center, Institute of Medical Science, University of Tokyo
 (partly supported by Science and Technology Agency).
 Location/Qualifiers
 1. 2830
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="HEP09821"
 /cell_line="HepG2"
 /cell_type="hepatoma"
 /clone_lib="HEP"
 /notes="Cloning vector: pME18SFL3"
 251..1537
 /notes="unnamed protein product"
 /codon_start=1
 /protein_id="BAB85033.1"
 /db_xref="GI:18676819"
 /translation="MGPPPGAGVSCGCGFSLRLAWCFLLALLSPOAGSRGAEAVWT
 AYLNVSRVPHTVNRTWELSEGVGGQSPLEPVAGVLVPPDGPALNCPHTNF
 TVPTVWSTQVSWLALIQGGCTPADKHLAYERGAGVIFNFPGRNREVI PMSH
 PGADVIAIMGLKTIQSTQIGIQTWIEVKGKHPWNHYSIPFVSVPFII
 TAATVGYFIFYSARLRNARQKORQLKADAKAIGRLQRLTKQDKETGPDDGS
 CAVCIELYFNDLVRLITCNHIFKTPDPLLEHRTCPMKCDILKALGIREDVEDG
 SVSLQVPVNEISNSASSHEEDNRSETASGVASVQGTDEPPLBEHVQSTNESLQLVN
 HEANSVAVDVIPHVDNPTFEDETPNQAVREIKS"

FEATURES

source

CDS

BASE COUNT

ORIGIN

Query Match 93.6%; Score 1659.6; DB 9; Length 2830;
 Best Local Similarity 99.2%; Pred. No. 0;
 Matches 1688; Conservative 0; Mismatches 10; Indels 3; Gaps 2;

QY 74 CCGAGGAGCTGCATCTCGGCAACCTGTGTGCTGAGCGCTACGTGCTCTCGGCTCCGACG 133
 DB 62 CCGAGGAGCTGCATCTCGGCAACCTGTGTGCTGAGCGCTACGTGCTCTCGGCTCCGACG 121
 QY 134 TAGCTCGAGCTCCCGAGCTCCTCACTTCCTTCCCGAGCTGCGCGGACCTGCTCAAG 193
 DB 122 TAGCTCGAGCTCCCGAGCTCCTCACTTCCTTCCCGAGCTGCGCGGACCTGCTCAAG 181
 QY 194 ACCAGGCTCTGCGCAAGCGCTAGGAGGCGCGTGCAGGCGCGCTAGGGAACCTGCGGAGC 253
 DB 182 ACCAGGCTCTGCGCAAGCGCTAGGAGGCGCGTGCAGGCGCGCTAGGGAACCTGCGGAGC 241
 QY 254 GCGCGCGCATGGGCGCGCGCTGCGGCGCGGCTCTCTGCGCGGCTGCTGCGGCTTT 313
 DB 242 GCGCGCGCATGGGCGCGCGCTGCGGCGCGGCTCTCTGCGCGGCTGCTGCGGCTTT 301
 QY 314 TCCAGATTGCTGGCATGTGCTTCTGCTGCGCTGAGTCCGAGGACCGCGGCTCCCGG 373
 DB 302 TCCAGATTGCTGGCATGTGCTTCTGCTGCGCTGAGTCCGAGGACCGCGGCTCCCGG 361
 QY 374 GGGCTGAAGCAGTGTGGACCGGCTACCTCAACGCTGCTGCGGCGGTTCCGCGACAGGGA 433
 DB 362 GGGCTGAAGCAGTGTGGACCGGCTACCTCAACGCTGCTGCGGCGGTTCCGCGACAGGGA 421
 QY 434 GTGAACCGTACGTGTGGAGCTGAGCGAGGCGGTGTACGCGCAGGACTCGCGCTG 493

Db 422 GTGAACCGTACGCTGTGGAGAGCTGAGCGAGGAGGCGGTGTACGCGCAGGACTCGCGCGTGTG 481
 QY 494 GAGCCTGTGGCTGGGCTCTGTGTACCGCCCGACGCGGCGCGGCGGCTTAAAGCCTGTGAAC 553
 Db 482 GAGCCTGTGGCTGGGCTCTGTGTACCGCCCGACGCGGCGCGGCGGCTTAAAGCCTGTGAAC 541
 QY 554 CGGACACGAAATTTACGGTGGCCACGGTTTGGGGAGACCGGTGCAAGTCTCTTGGTGTG 613
 Db 542 CGGACACGAAATTTACGGTGGCCACGGTTTGGGGAGACCGGTGCAAGTCTCTTGGTGTG 601
 QY 614 GGCCTCATCCAAACGCGCGGCGGCTGACCTTTCGACAGCAAGATCCATCTCGCTTATGAG 673
 Db 602 GGCCTCATCCAAACGCGCGGCGGCTGACCTTTCGACAGCAAGATCCATCTCGCTTATGAG 661
 QY 674 AGATGGGCGTCTGGAGCGCTCATCTTTAACTTTCGCGGACCCGCAATGAGGTGATCCCC 733
 Db 662 AGAGGGGCGTCTGGAGCGCTCATCTTTAACTTTCGCGGACCCGCAATGAGGTGATCCCC 721
 QY 734 ATGTCTCACCGGCTGCGAGTACATTTGTCGATCATGTCGCAATCTGAAAGGACACA 793
 Db 722 ATGTCTCACCGGCTGCGAGTACATTTGTCGATCATGTCGCAATCTGAAAGGACACA 781
 QY 794 AAAATTTCTGCAATCTTATTTCAAAGGCGCATACAAGTGACAATGGTGCATAGAAAGTAGGAAA 853
 Db 782 AAAATTTCTGCAATCTTATTTCAAAGGCGCATACAAGTGACAATGGTGCATAGAAAGTAGGAAA 841
 QY 854 AAAATGGCCCTTGGGTGAATCACTATTTCAATTTTTTTCGTTCTGTGTCCTTTTTTATT 913
 Db 842 AAAATGGCCCTTGGGTGAATCACTATTTCAATTTTTTTCGTTCTGTGTCCTTTTTTATT 901
 QY 914 ATTACGGGCGCAACTGTGGGCTATTTTATCTTTTCTGCTCGAAGCTACGGAATGCA 973
 Db 902 ATTACGGGCGCAACTGTGGGCTATTTTATCTTTTCTGCTCGAAGCTACGGAATGCA 961
 QY 974 AGAGCTCAAAGCAGGAGCAGAGGCAATTAAGAGCAGATGCTAAAGAAAGCTATTGGAAAG 1033
 Db 962 AGAGCTCAAAGCAGGAGCAGAGGCAATTAAGAGCAGATGCTAAAGAAAGCTATTGGAAAG 1021
 QY 1034 CTTCAACTACGACACACTGAAACAAGGAGACAAGGAAATTTGGCCCTGTATGGAGATAGTTGT 1093
 Db 1022 CTTCAACTACGACACACTGAAACAAGGAGACAAGGAAATTTGGCCCTGTATGGAGATAGTTGT 1081
 QY 1094 GCTGTGTCGCTTGAATTTGATTAACCAATGATTTGCTACGATCTTAAAGCTGCAACCAT 1153
 Db 1082 GCTGTGTCGCTTGAATTTGATTAACCAATGATTTGCTACGATCTTAAAGCTGCAACCAT 1141
 QY 1154 ATTTTCCATAAGACATGTGTGACCCATGGCTGTAAAAACAAGAGCTTGGCCCATGTGC 1213
 Db 1142 ATTTTCCATAAGACATGTGTGACCCATGGCTGTAAAAACAAGAGCTTGGCCCATGTGC 1201
 QY 1214 AAATGTGACATCTCAAGAGCTTTGGGAATTTGAGGTGAGATTTGAAGATGGAATCAGTGCT 1273
 Db 1202 AAATGTGACATCTCAAGAGCTTTGGGAATTTGAGGTGAGATTTGAAGATGGAATCAGTGCT 1261
 QY 1274 TTACAAGTCCCTGTATCCAAATGAAATATCTAATAGTGCCTCTCCCTGAGAGGATAT 1333
 Db 1262 TTACAAGTCCCTGTATCCAAATGAAATATCTAATAGTGCCTCTCCCTGAGAGGATAT 1321
 QY 1334 GCGAGCGAGACCGCATCATCTGGATATGCTTCAGTACAGGGAACAGATGAAACCGCCTCTG 1393
 Db 1322 GCGAGCGAGACCGCATCATCTGGATATGCTTCAGTACAGGGAACAGATGAAACCGCCTCTG 1381
 QY 1394 GAGGAAACGCTGAGTCAACAAATGAAAGTCTACAGCTGGTAAACCATGAAAGCAAAATCT 1453
 Db 1382 GAGGAAACGCTGAGTCAACAAATGAAAGTCTACAGCTGGTAAACCATGAAAGCAAAATCT 1441
 QY 1454 GTGGCAGTGGATGTTATTCCTCATGTGACAAACCAACCTTTGAGAGAGAGCAAACTCCT 1513
 Db 1442 GTGGCAGTGGATGTTATTCCTCATGTGACAAACCAACCTTTGAGAGAGAGCAAACTCCT 1501
 QY 1514 AATCAAGAGACTGCTGTTTCGAGAAATTAATCTTTAAATCTGTGTAAATGAAACTTGA 1573
 Db 1502 AATCAAGAGACTGCTGTTTCGAGAAATTAATCTTTAAATCTGTGTAAATGAAACTTGA 1561

Db 1141 GGAGAACACGTCAGTCAACAAATGAAAGTCTACAGCTGGTAAACCATGAACCAATTC 1200

Qy 1453 TGTGGCAGTGAATTTATCTCTCATGTCAGAACCCACCTTTGAAGAACCAACTCC 1512

Db 1201 TGTGGCAGTGAATTTATCTCTCATGTCAGAACCCACCTTTGAAGAACCAACTCC 1260

Qy 1513 TAATCAAGAGACTCTGTTGAGAGAAATTAATCTTAA 1549

Db 1261 TAATCAAGAGACTCTGTTGAGAGAAATTAATCTTAA 1297

RESULT 4

BC010477

LOCUS

DEFINITION Mus musculus, RIKEN cDNA 1300002C13 gene, clone MGC:19395

ACCESSION BC010477

VERSION

KEYWORDS

SOURCE MGC.

ORGANISM

Mus musculus

house mouse.

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 1928)

Strausberg.R.

Direct Submission

Submitted (10-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgaps@mail.nih.gov

Tissue Procurement: Jothar Hennighausen Ph.D., Chu-Xia Deng Ph.D. cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: amg@bcm.tmc.edu

Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 23 Row: b Column: 12

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 12963610.

FEATURES

source

1..1928

/organism="Mus musculus"

/db_xref="taxon:10090"

/map="129,C57BL/6J,FVB/N"

/clone="MGC:19395 IMAGE:3154433"

/tissue types="Mammary tumor. Brca1-/-fl; MMTV-Cre model. 10 months old, gross tissue."

/clone_lib="NCI CGAP_Mam3"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

176..1462

/codon_start=1

/product="RIKEN cDNA 1300002C13 gene"

/protein_id="AAH10477.1"

/db_xref="GI:14714673"

/translation="MGPPPGIGVYCRGGCAARLLAWCFLLASPHAPSGRAEAVVT AYLNWNRVPHTCVNTVWELSEGVYGDSPLEPVSGVLVDPDGPALNACNPTNF TPTVWGSTVQVSWLALIQRGSGCTFADKHLASERGASGAVIFPPGTRNRIEMSH PGAGDVAIMIGKGTILQSQVQVTVIEVQKGGPWNHYISFFVSVSFFII TAATVTFYFSARLNRARQSRKQRLKAKKAIGKLQRLTKQDGPETGPDGS CAVCIELYFPNDILVILTCNHIHFHKTCDPFWLLEHRTCPMKCKDILKALGIEVDVEDG

CDS

SVSLQVPVNEASNTASPHEDSRSETASSGYASVQGADEPPLEEHAQSANENLQLVN
HEANSVADVPHVDNPFTFEEDETPQEAARVRIKS"

BASE COUNT 482 a 466 c 498 g 482 t

ORIGIN

Query Match 67.9%; Score 1205.2; DB 10; Length 1928;
Best Local Similarity 89.6%; Pred. No. 1.4e-265;
Matches 1364; Conservative 0; Mismatches 149; Indels 10; Gaps 6;

Qy 252 GCGCGCGCCATGGGGCGCGCGCTGGGGCGGGGCTCTCTGCGCGCGGTGGCTGCGGCT 311

Db 165 GAGTGGCGCCATGGGGCGCGCGCTGGGGCGGGGCTCTCTGCGCGCGGTGGCTGCGGAG 224

Qy 312 TTTCCAGATTTGCTGGCATGGTCTTCTGCTGGCCCTGAGTCCGAGGACACCCGGTTCCC 371

Db 225 CTGCCCGGCTACTGGCTTTGGTCTTCTCTTCTGCTCTGAGTCCGACGCGCGCGGTCCC 284

Qy 372 GGGGGGCTGAAGACAGTGTGGACCGGTACTCAAGCTGCTCTGGCGGGTCCGACACGG 431

Db 285 GGGGAGCGAAGCGGTGTGGACCGGTACTCAAGCTGCTCTGGCGGGTCCGACACCG 344

Qy 432 GAGTGAACCGTACCGTGTGGAGCTGAGCGAGGAGGGGCTGTACGGCCAGGACTCGCGC 491

Db 345 GAGTGAACCGCACCGGTGTGGAGCTGAGCGAGGAGGGGCTGTACGGCCAGGACTCGCGC 404

Qy 492 TGGAGCTGTGGCTGGGGTCTCTGTTACCGCCGAGCGGGCCCGGGCGCTTAACGCTGTA 551

Db 405 TGGAGCGGCTCTCGGGGCTCTGTTACCGCCGAGCGGGCCCGGGCGCTTAACGCTGTA 464

Qy 552 ACCCGCACACCAATTTTCCAGGTGCCCGCTTTGGGGAGGACGCGTCAAGTCTCTTGGT 611

Db 465 ACCCGCACACCAATTTTCCAGGTGCCCGCTTTGGGGAGGACGCGTCAAGTCTCTTGGT 524

Qy 612 TGGCCCTCATCAACCGCGCGGGGCTGCACCTTCGCGACAAAGATCCATCTGGCTTATG 671

Db 525 TGGCCCTCATCAACCGCGGTGGAGGCTGCACCTTCGCGGACAAAGATCCATCTGGCTTATG 584

Qy 672 AGAGATGGGCTCTGGAGCGCTCATCTTTAACTTCCCGGAGACCGGAAATGAGTCACTCC 731

Db 585 AGAGAGGGGCTCTGGAGCGCTCATCTTTAACTTCCCGGAGACCGGAAATGAGTCACTCC 644

Qy 732 CCATCTCTCACCCGGGTTCAGTACATCTTTGCAATCATGATGGCAATCTGAAAGGCA 791

Db 645 CCATCTCTCACCCGGGTTCAGTACATCTTTGCAATCATGATGGCAATCTGAAAGGCA 704

Qy 792 CAAAAATTTCTCAATTTCAAGAGGATCAAGTCAATGGTCAATAGAGTAGGGA 851

Db 705 CAAAAATTTCTCAATTTCAAGAGGATCAAGTCAATGGTCAATAGAGTAGGGA 764

Qy 852 AAAAAATTTCTCAATTTCAAGAGGATCAAGTCAATGGTCAATAGAGTAGGGA 911

Db 765 AAAAAATTTCTCAATTTCAAGAGGATCAAGTCAATGGTCAATAGAGTAGGGA 824

Qy 912 TTAATTAACCGCGCAATCTGGGCTATTTTATCTTTTATCTGCTCGAAGGCTACGGAATG 971

Db 825 TTAATTAACCGCGCAATCTGGGCTATTTTATCTTTTATCTGCTCGAAGATACGAATG 884

Qy 972 CAAAGCTCAAGAGGAGGAGGCAATTTAAAGGAGAGTGTCTTAAAGAGTATTTGAA 1031

Db 885 CAAAGCTCAAGAGGAGGAGGCAATTTAAAGGAGAGTGTCTTAAAGAGTATTTGAA 944

Qy 1032 GGCCTCAACTAGCACACTGAAACAAAGAGGAGGAGGAAATTTGGCCCTGATGGAGTAGTT 1091

Db 945 AGCTTCAGCTGGCACCTTTGAAACAAAGAGGAGGAGGAAATTTGGCCCTGATGGAGTAGTT 1004

Qy 1092 GTGCTGTGTGCAATTTGAAATTTGTAACCAAAATGATTTGGTACGCAATCTTAAACGTGCAAC 1151

Db 1005 GTGCTGTGTGCAATTTGAGCTCTTAAAGCAAAATGATTTGGTGGGATCTCACTACCTGCAATC 1064

Qy 1152 ATATTTTCCATAAGACATGTGTTGACCACTGGCTGTTAAAAACAAGAATTTGCCCATGT 1211

Db 1065 ATATTTTCCATAAGACATGTGTTGACCACTGGCTGTTAAAAACAAGAATTTGCCCATGT 1124

QY 792 CAAAAATTCGCAATCTATTCAAAGGCGCATACAAAGTGAATGGTGCATAGAGTAGGGA 851
 Db 716 CAAAAATTCGCAATCTATTCAAAGGCGCATACAAAGTGAATGGTGCATAGAGTAGGGA 775
 QY 852 AAAAAATCGCCCTCGGCTGAATCACTATTCAATTTTTTCTGTTCTGTTCTTTT 911
 Db 776 AAAAAATCGCCCTCGGCTGAATCACTATTCAATTTTTTCTGTTCTGTTCTTTTCA 835
 QY 912 TTATTAACGGCGCAACTGGGCTATTTTATCTTTTATCTCTCGAAGGCTACGAATG 971
 Db 836 TAAATACGGCAGCAACCGTGGGCTATTTTATCTTTTATCTCTCGAAGATTAACGAATG 895
 QY 972 CAAGAGCTCAAGAGCAGAGAGGCAATTAAGCGCATGCTTAAAAAGCTATTGAA 1031
 Db 896 CAAGAGCTCAAGAGCAGAGAGGCAATTAAGCGCATGCTTAAAAAGCTATTGAA 955
 QY 1032 GGCTTCAACTAGCAGCACTGAAACAGGAGACAAAGAAATTTGGCCCTGATGAGATAGT 1091
 Db 956 AGCTTCAGCTGGCAGCTTGAACAGGAGACAAAGAAATTTGGCCCTGATGAGATAGT 1015
 QY 1092 GTGCTGTGTCATTAATGTTATTAACAAATGATTTGGTACGATCTTAACGTCAACC 1151
 Db 1016 GTGCTGTGTCATTAATGTTATTAACAAATGATTTGGTACGATCTTAACGTCAATC 1075
 QY 1152 ATATTTTCCATAGACATGTTGACCCATGGCTGTTAAAAACAAAGACTTGCCTCATGT 1211
 Db 1076 ATATTTTCCATAGACATGTTGACCCATGGCTGTTAAAAACAAAGACTTGCCTCATGT 1135
 QY 1212 GCAATATGTGACATCTCAAGAGCTTTGGGAATTTGAGTGTGATGTTGAAGATGATCAGTGT 1271
 Db 1136 GCAAGTGTGACATCTCAAGAGCTTTGGGAATTTGAGTGTGATGTTGAAGATGATCAGTGT 1195
 QY 1272 CTTTAAAGTCCCTGATCAATGAATATCTAATAGTTCCTTCCATGAAGAGGATA 1331
 Db 1196 CTTTAAAGTTCCTGTTCTAATGAAGCATCTAATAGTTCCTTCCATGAAGAGGATA 1255
 QY 1332 ATGCGAGCAGAGCCGATCATCTGGATATGCTTCAGTACAGGAGACAGATGAACGCTC 1391
 Db 1256 GTCCAGTGAAGTGCATCATCTGGATATGCTTCAGTACAGGAGACAGATGAACGCTC 1315
 QY 1392 TGGAGGAACACGTCGAGTCAACAAATGAAAGTCTACAGCTGGTAAACCATGAAGCAAAAT 1451
 Db 1316 TGGAGGAACATGCGCAGTCAGCAATGAAATCTACAGCTGGTAAACCATGAAGCAAAAT 1375
 QY 1452 CTGTGCGATGATGTTATCTCATGTGNGACAACCAACCTTTGAAGAGAGCAATC 1511
 Db 1376 CTGTGCGCGTGGATGTTGTCGCCATGTTGACAAACCAACCTTTGAAGAGAGCAATC 1435
 QY 1512 CTAATCAAGAGACTGCTGTTGAGAAATTTAAATCTTT-AAAACTGTGTAATAGAAACT 1570
 Db 1436 CTGATCAAGAGGAGCTGTTGCGGAGATTAATCTTTAAATATCTGTGTCATAGAAACT 1495
 QY 1571 TGAACCATTAG-TAATAACAGAACTGCCAATCAGGCGCTAGTTT-CTATTAATAAATGG 1628
 Db 1496 TGAACCGTTAGTTAACAACAGGACTGCCAATCAGGCGCTAGTTTACTATGAATGAATGG 1555
 QY 1629 ATAAATTTAATAAATAAGAGTACTCAAAAGTGTCTAGATGACTAATATTAATCTATA 1688
 Db 1556 GTAAA--CGTAAAAAAGAAATGATGCTGAAAGTGTGAGGTAACCTTATATTAATATA 1612
 QY 1689 GTTAAATGGCTTTAAAAATATTAACTGCTTAACCTTTTTCACCAAACTCATATTATA 1748
 Db 1613 GTTAAA--TGGCTTAACATATTATACCCAGTACCGTTT--CCAAACTTCACTAAG 1668
 QY 1749 TTTTTCATAGGCAAGTTTCTCT 1771
 Db 1669 TTTTTCATAGGCAAGTTTCTCT 1691

RESULT 6
 AF426411
 LOCUS

1287 bp mRNA linear ROD 22-NOV-2001

DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

Mus musculus zinc ring finger-containing protein GRAIL mRNA,
 complete cds.
 AF426411
 AF426411.1 GI:17046405
 Mus musculus.
 Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 1 (bases 1 to 1287)
 FORD, G.S., BLOOM, D., PARAGAE, V., ANANDASABAPATHY, N., SKRENTA, H.,
 ERMANN, J., CRON, R.Q., LEWIS, D.B. and FATHMAN, G.
 GRAIL: A novel gene expressed in anergic T cells that inhibits
 NFAT/AP-1 activation and IL-2 transcription
 JOURNAL
 REFERENCE
 2 (bases 1 to 1287)
 FORD, G.S., BLOOM, D., PARAGAE, V., ANANDASABAPATHY, N., SKRENTA, H.,
 ERMANN, J., CRON, R.Q., LEWIS, D.B. and FATHMAN, G.
 Direct Submission
 Submitted (01-OCT-2001) Medicine, Immunology and Rheumatology,
 Stanford University School of Medicine, 300 Pasteur Drive,
 Stanford, CA 94305, USA

FEATURES
 source

1..1287
 /organism="Mus musculus"
 /strain="DBA/2"
 /db_xref="taxon:10090"
 /clone="11.3.7"
 /cell_type="anergic T cell"
 1..1287
 /codon_start=1
 /product="zinc ring finger-containing protein GRAIL"
 /protein_id="AAL34514.1"
 /db_xref="GI:17046406"

CDS

BASE COUNT 308 a 320 c 359 g 300 t
 ORIGIN
 Query Match 62.4%; Score 1106.8; DB 10; Length 1287;
 Best Local Similarity 91.2%; Pred. No. 4.7e-243;
 Matches 1174; Conservative 0; Mismatches 113; Indels 0; Gaps 0;
 QY 263 ATGGGGCGCGCGCTGGGGCGGGGTCTCTGCGGGGTGGCTGCGGCTTTCCAGATTG 322
 Db 1 ATGGGGCGCGCGCGCGGGATCGGGGTCTACTGCGGGGTGGCTGCGGAGCTGCCCGGCTA 60
 QY 323 CTGGCATGGTCTCTCTGCTGGCCCTGAGTCCGAGCACCCGGTTCCTCCGGGGGCTGAA 382
 Db 61 CTGGCTTGGTCTCTCTCTGCTGCTGAGTCCGACGCGCCCGTTCCTCCGGGAGCGGAA 120
 QY 383 GCAGTGTGGACCGGTACCTCAACGTGTCTGCGGGTTCGACACGCGGAGTGAACCGT 442
 Db 121 GCGGTGTGGACCGGTACCTCAACGTGTCTGCGGGTTCGACACACCGGAGTGAACCGC 180
 QY 443 ACAGTGTGGAGCTGAGCAGGAGGGGTGTACGGCAGAGCTCGCGGCTGGAGCCTGTG 502
 Db 181 ACAGTGTGGAGCTGAGCAGGAGGGGTGTACGGCAGAGCTCGCGGCTGGAGCCTGTG 240
 QY 503 GCTGGGGTCTCTGGTACCGCCCGAGCGGGCGCGCTTAACGCCCTGTAAACCGGCACACG 562
 Db 241 TCGGGGTCTCTGGTACCGCCCGAGCGGGCGCGCTCAACGCCCTGTAAACCGGCACAC 300
 QY 563 AATTTCAGGTGCGCCACGGTTTGGGGAAGACACCGTGAAGTCTCTTGGTTGGCCCTCATC 622
 Db 301 AATTTCAGGTGCGCCACGGTTTGGGGAAGACACCGTGAAGTCTCTTGGTTGGCCCTCATC 360

BASE COUNT	670 a	277 c	321 g	603 t	ORIGIN
Query Match	40.3%	Score 714.4	DB 9	Length 1871	
Best Local Similarity	97.9%	Pred. No. 5e-153			
Matches 755; Conservative	0	Mismatches 12	Indels 4	Gaps 3	
QY 1005	AGGCAGATGCTAAAAAGCTATTGGAGGCTTCAACTACGCACACTGAAACAAGGAGACA	1064			
DB 1	AGGCAGATGCTAAAAAGCGTTAGGAAGGCTTCAACTACTCACACTGAAACAAGGAGACA	60			
QY 1065	AGGAATTTGGCCCTGATGGAGATAGTTGTCTGTGTGCATTGTAATTTGATATAAACCAATG	1124			
DB 61	AGGAATTTGGCCCTGATGGAGATAGTTGTCTGTGTGCATTGTAATTTGATATAAACCAATG	120			
QY 1125	ATTTTGGTAC-GCATCTTTAAGCTGCACCACTATTTTCCATAAGACATGTGTTGACCCCATGG	1183			
DB 121	ATTTTGGTACAGCATCTTTAAGTGCACCACTATTTTCCATAAGACATGTGTTGACCCCATGG	180			
QY 1184	CTGTTAAAAACAAGACTTTGCCCCCATGTGCAAAATGTGACATACTCAAAAGCTTTGGGGAATT	1243			
DB 181	CTGTTAGACACAGAGACTTTGCCCCATGTGCAATGTGACATACTCAAAAGCTTTGGGGAATT	240			
QY 1244	GAGGTGGATGTTGAAGATGGATCAGTGTCTTTTCAAGTCCCTGTATCCAAATGAAATATCT	1303			
DB 241	GAGGTGGATGTTGAAGATGGATCAGCGTCTTTTCAAGTCCCTGTATCCAAATGAAATATCT	300			
QY 1304	AATAGTGCCTCCTCCCATGAGAGGATTAATCGCAGCGAGACCCGATCATCTGGATATGCT	1363			
DB 301	AATAGTGCCTCCTCCCATGAGAGGATTAATCGCAGCGAGACCCGATCATCTGGATATGCT	360			

Db	181	CTGTTAGAACACAGGACTTGTGCCCATGTGTGCAAAATGTGCATATCTCAAAGCTTTGGGAATT	240
QY	1244	GAGGTGGATGTTTGAAGATGGATCAGTGTCTTTTACAAGTCCCTGTATCCCATGAATATCT	1303
Db	241	GAGTGGATGTTTGAAGATGGATCAGCGTCTTTACAAGTCCCTGTATCCCATGAATATCT	300
QY	1304	AATAGTGCTCTCCCATGAAGAGGATATCGCAGCGAGACCGCATCATCTGGATATGCT	1363
Db	301	AATAGTGCCTCTCCCATGAAGAGGATATCGCAGCGAGACCGCATCATCTGGATATGCT	360

QY 1364 TCAGTACAGGAAACAGATGAACCGCCTCTGGAGGAAACAGTGCAGTCAACAAATGAAAGT 1423
 |||||
 Db 361 TCAGTACAGGAAACAGATGAACCGCCTCTGGAGGAAACAGTGCAGTCAACAAATGAAAGT 420
 |||||
 QY 1424 CTACAGCTGGTAAACCATCAAGCAAAATCTGTGGCAGTGGATGTTATTCCTCATGTGAC 1483
 |||||
 Db 421 CTACAGCTGGTAAACCATCAAGCAAAATCTGTGGCAGTGGATGTTATTCCTCATGTGAC 480
 |||||
 QY 1484 AACCCAAACCTTTGAAGAACAGAAATCTCTAATCAAGAGACTGCTGTTTCGAGAAATTA 1543
 |||||
 Db 481 AACCCAAACCTTTGAAGAACAGAAATCTCTAATCAAGAGACTGCTGTTTCGAGAAATTA 540
 |||||
 QY 1544 TCTTAAATCTGTGTAATAGAAACTTGAACCATAGTAATAACAGAACTGCCAATCAG 1603
 |||||
 Db 541 TCTTAAATCTGTGTAATAGAAACTTGAACCATAGTAATAACAGAACTGCCAATCAG 600
 |||||
 QY 1604 GGCCTAGTTCTTATTAATAATTTGATAATTTAATAAATAAGAGTGTACTGAAAGTG 1663
 |||||
 Db 601 GGCCTAGTTCTTATTAATAATTTGATAATTTAATAAATAAGAGTGTACTGAAAGTG 660
 |||||
 QY 1664 CTCAGATGACTAATATATGCTATAGTTAAATGGCTTAAATAATTTAACTGTTAACT 1723
 |||||
 Db 661 CTCAGATGACTAATATATGCTATAGTTAAATGGCTTAAATAATTTAACTGTTAACT 718
 |||||
 QY 1724 TTTTCCCAAACTCATTAATATTTTTCATAGCAAGTTCTCTCTCAG 1774
 |||||
 Db 719 TTTTCCCA-AAAACTCATTAATATTTTTCATAGCAAGTTCTCTCTCAG 768
 |||||

RESULT 8
 AL391315 Human DNA sequence from clone RP11-150F24 on chromosome X, complete sequence.
 LOCUS AL391315
 DEFINITION AL391315.22 GI:14575271
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE HTG.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 169812)
 Brown, A.
 Direct Submission
 Submitted (27-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 requests: clonerequest@sanger.ac.uk
 On Jun 28, 2001 this sequence version replaced gi:14455902.
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; SW, SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/ChrX
 RP11-150F24 is from the library RPCI-11.1 constructed by the group of Pieter de Jong. For further details see

http://www.chori.org/bacpac/home.htm

VECTOR: pBACes.6
 This sequence is the entire insert of clone RP11-150F24. The true left end of clone RP11-321G1 is at 6861 in this sequence. The true right end of clone RP11-697G3 is at 161410 in this sequence.

FEATURES

source

Location/Qualifiers

1..169812

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="X"

/clone="RP11-150F24"

/clone_lib="RPCI-11.1"

12..569

/note="match: GSS: Em:AQ375615"

/note="2 copies 36 mer 90% conserved"

repeat_region

512..729

/note="MIR repeat: matches 3..238 of consensus"

repeat_region

1338..1638

/note="AluY repeat: matches 1..304 of consensus"

repeat_region

3599..3642

/note="L2 repeat: matches 2707..2750 of consensus"

repeat_region

3722..3899

/note="MIR repeat: matches 12..171 of consensus"

repeat_region

3915..4028

/note="MIR repeat: matches 131..238 of consensus"

repeat_region

5121..5197

/note="L2 repeat: matches 2611..2710 of consensus"

repeat_region

6734..7029

/note="AluSp repeat: matches 1..297 of consensus"

repeat_region

7660..7955

/note="AluSq repeat: matches 1..308 of consensus"

repeat_region

8891..8940

/note="L1P9 repeat: matches 6114..6163 of consensus"

repeat_region

10067..10370

/note="AluX repeat: matches 1..297 of consensus"

repeat_region

10411..10695

/note="AluX repeat: matches 1..288 of consensus"

repeat_region

10737..10879

/note="MER3 repeat: matches 5..154 of consensus"

repeat_region

11078..11380

/note="AluY repeat: matches 1..300 of consensus"

repeat_region

11571..12560

/note="L1MD2 repeat: matches 5328..6331 of consensus"

repeat_region

13476..13794

/note="L1ME1 repeat: matches 5837..6165 of consensus"

repeat_region

13906..14041

/note="L2 repeat: matches 2536..2671 of consensus"

repeat_region

14106..14411

/note="AluY repeat: matches 2..306 of consensus"

repeat_region

16413..16704

/note="AluX repeat: matches 1..295 of consensus"

repeat_region

17566..17697

/note="FLAM_C repeat: matches 1..132 of consensus"

repeat_region

17839..18080

/note="MIR repeat: matches 3..260 of consensus"

repeat_region

18961..19153

/note="MIR repeat: matches 30..262 of consensus"

repeat_region

19562..19658

/note="MER5A repeat: matches 12..109 of consensus"

repeat_region

20326..20413

/note="L1MC4 repeat: matches 7861..7948 of consensus"

repeat_region

20471..20754

/note="AluSq repeat: matches 1..286 of consensus"

repeat_region

20755..20782

/note="14 copies 2 mer ta 89% conserved"

repeat_region

20999..21184

/note="L2 repeat: matches 2529..2729 of consensus"

repeat_region

21194..21952

/note="L2 repeat: matches 1730..2546 of consensus"

repeat_region

22817..23035

/note="MIR repeat: matches 15..238 of consensus"

repeat_region

23641..24552

Misc_feature


```

VERSION      AL591467.3   GI:14348539
KEYWORDS     HTG; HTGS_PHASE2; HTGS_CANCELLED.
SOURCE       human.
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 169863)
AUTHORS      Bird.C. Submission
TITLE        Direct Submission
JOURNAL      Submitted (10-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
             CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
COMMENT      requests: clonerequest@sanger.ac.uk
             On Jun 12, 2001 this sequence version replaced gi:14268300.
             ----- Genome Center
             Center: Sanger Centre
             Center code: SC
             Web site: http://www.sanger.ac.uk
             Contact: humquery@sanger.ac.uk
             ----- Project Information
             Center project name: BA697G3
             ----- Summary Statistics
             Sequencing program: XGAP4; version 4.5
             Chemistry: Dye-terminator Big Dye; 100% of reads
             Consensus quality: 157318 bases at least Q40
             Consensus quality: 157886 bases at least Q30
             Consensus quality: 158243 bases at least Q20
             Insert size: 169863; sum-of-contigs
             Insert size: 169949; 0.8% error; agarose-fp
             Quality coverage: 5.85x in Q20 bases; sum-of-contigs Quality
             coverage: 6.21x in Q20 bases; agarose-fp
             -----
             * NOTE: This is a 'working draft' sequence.
             * This sequence will be replaced
             * by the finished sequence as soon as it is available and
             * the accession number will be preserved.

FEATURES             Location/Qualifiers
     source           1..169863
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /clone="RP11-697G3"
                     /clone_lib="RPC1-11.3"
     misc_feature     1..169863
                     /note="assembly fragment: 02868"
BASE COUNT          51891 a 31100 c 31417 g 55455 t
ORIGIN
Query Match        38.0%; Score 674.4; DB 2; Length 169863;
Best Local Similarity 99.9%; Pred. No. 1.9e-143;
Matches 675; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 74 CCGAGAGCTGATCTGGGCAACCTGTGTCTGACGCTAGCTGCTCTCTGCTCCGACG 133
Db 23909 CCGAGAGCTGATCTGGGCAACCTGTGTCTGACGCTAGCTGCTCTCTGCTCCGACG 23968
QY 134 TAGCTCGCAGCTCCCGAGTCTCACTCCATTCCTTCCCACTGGCGGCACTGCTCAAG 193
Db 23969 TAGCTCGCAGCTCCCGAGTCTCACTCCATTCCTTCCCACTGGCGGCACTGCTCAAG 24028
QY 194 ACCAGGGTCTTGCCAAAGCGCTAGGAGGCGCGTGCAGGGGCGCTAGGAACTCGCGAGC 253
Db 24029 ACCAGGGTCTTGCCAAAGCGCTAGGAGGCGCGTGCAGGGGCGCTAGGAACTCGCGAGC 24088
QY 254 GGGCGGCGATGGGCGCGCGCTGGGCGCGGGTCTCTCTCCCGGGTGGCTGGCGTTT 313
Db 24089 GGGCGGCGATGGGCGCGCGCTGGGCGCGGGTCTCTCTCCCGGGTGGCTGGCGTTT 24148
QY 314 TCCAGATTGCTGGCATGGTGTCTTCTGCTGGCCCTGAGTCCGAGGCAACCGGTTCCCGG 373
Db 24149 TCCAGATTGCTGGCATGGTGTCTTCTGCTGGCCCTGAGTCCGAGGCAACCGGTTCCCGG 24208
QY 374 GGGGCTGAAGCATGTGTGACCGCGTACCTCAACGTGTCTCTGGCGGGTTCCGCAACACGGA 433

```

```

Db 24209 GGGGCTGAAGCAGTGTGACCGCGTACCTCAACGTGTCTTGGCGGGTTCCGACACGGA 24268
QY 434 GTGAACCGTACCGTGTGGGAGCTGAGCGAGGAGGCGGTGTACGGCCAGGACTCGCGCTG 493
Db 24269 GTGAACCGTACCGTGTGGGAGCTGAGCGAGGAGGCGGTGTACGGCCAGGACTCGCGCTG 24328
QY 494 GAGCCTGTGGTGGGCTCTGTGTACCGCCGACCGGCGCGGCGCTTAAAGCCTGTAAAC 553
Db 24329 GAGCCTGTGGTGGGCTCTGTGTACCGCCGACCGGCGCGGCGCTTAAAGCCTGTAAAC 24388
QY 554 CGGCACACGAATTTTCCACGGTGTCCACCGTGTGGGGAAGCACCGTGCAAGTCTCTTGGTGT 613
Db 24389 CGGCACACGAATTTTCCACGGTGTCCACCGTGTGGGGAAGCACCGTGCAAGTCTCTTGGTGT 24448
QY 614 GCCCTCATCCAAACCGCGCGGGGTGTGACCTTTCGACAGCAAGATCCATCTCTGGCTTATGAG 673
Db 24449 GCCCTCATCCAAACCGCGCGGGGTGTGACCTTTCGACAGCAAGATCCATCTCTGGCTTATGAG 24508
QY 674 AGATGGGGTCTGGAGCGGTCTATCTTTAACTTCCCGGAGCCGCAATGAGGTATCCGCC 733
Db 24509 AGAGGGGGTCTGGAGCGGTCTATCTTTAACTTCCCGGAGCCGCAATGAGGTATCCGCC 24568
QY 734 ATGTCTCACCCCGGGTG 749
Db 24569 ATGTCTCACCCCGGGTG 24584

RESULT 10
LOCUS       AX071980               404 bp    DNA    linear    PAT 25-JAN-2001
DEFINITION Sequence 2452 from Patent WO0102568.
ACCESSION  AX071980
VERSION     AX071980.1   GI:12582331
KEYWORDS   human.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 404)
AUTHORS     Williams,D.T., Escobedo,J., Innis,M.A., Garcia,P.D., Klinger,J.,
            Kassam,A., Reinhard,C., Randazzo,F., Kennedy,G.C., Pot,D.,
            Lamson,G., Drmanac,R., Crkenjakov,R., Drmanac,S., Dickson,M.,
            Labat I., Lezhkowitz,D., Kita,D., Garcia,V. and Strache-Crain,B.
            Human genes and gene expression products
            Patent: WO 0102568-A 2452 11-JAN-2001;
            CHIRON CORPORATION (US) ; HYSEQ, INC. (US)
FEATURES    Location/Qualifiers
     source           1..404
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT          134 a 81 c 90 g 99 t
ORIGIN
Query Match        22.1%; Score 392.2; DB 6; Length 404;
Best Local Similarity 99.0%; Pred. No. 2.8e-79;
Matches 394; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1194 ACAAGACTTGGCCCATGTGCAAAATGTGACATCTCAAAAGCTTGGGAATTGAGTGGATG 1253
Db 7 AGAGGACTTGGCCCATGTGCAAAATGTGACATCTCAAAAGCTTGGGAATTGAGTGGATG 66
QY 1254 TTGAAGATGATCAGTGTCTTTTACAAGTCCCTGTATCCAAATATCTTAATAGTGCCT 1313
Db 67 TTGAAGATGATCAGTGTCTTTTACAAGTCCCTGTATCCAAATATCTTAATAGTGCCT 126
QY 1314 CTTCCCATGAAGAGGATATTCGACGAGACCGCATCATCTGGATATGCTTCAGTACAGG 1373
Db 127 CTTCCCATGAAGAGGATATTCGACGAGACCGCATCATCTGGATATGCTTCAGTACAGG 186
QY 1374 GAACAGATGAACCGGCTCTGGAGGAACACGTCAGTCAACAAATGAAAGTCTACAGCTGG 1433
Db 187 GAACAGATGAACCGGCTCTGGAGGAACACGTCAGTCAACAAATGAAAGTCTACAGCTGG 246

```

QY 1434 TAAACCATGAAGCAAAATTCGTGGCAGTGGATGTTATTCCTCATGTNGACAACCCAACT 1493
 Db 247 TAAACCATGAAGCATATTCGTGGCAGTGGATGTTATTCCTCATGTNGACAACCCAACT 306
 QY 1494 TTGAAGAGAGCAAACTCCCTAATCAAGAGACTGCTGTTGAGAAATTAATCTTAAATC 1553
 Db 307 TTGAAGAGAGCAAACTCCCTAATCAAGAGACTGCTGTTGAGAAATTAATCTTAAATC 366
 QY 1554 TGTGTAATAGAAAATCTTGAACCATTTAGTAATAACAGA 1591
 Db 367 TGTGTAATAGAAAATCTTGAACCATTTAGTAATAACAGA 404
 RESULT 11
 AX341568
 LOCUS
 DEFINITION Sequence 1815 from Patent WO0196388. PAT 10-JAN-2002
 ACCESSION AX341568
 VERSION AX341568.1 GI:18137550
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Jiang, Y., Harlocker, S.L. and Secrist, H.
 TITLE Compositions and methods for the therapy and diagnosis of colon cancer
 JOURNAL Patent: WO 0196388-A 1815 20-DEC-2001;
 CORIXA CORPORATION (US)
 FEATURES
 source Location/Qualifiers
 1..646
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 226 a 103 c 106 g 208 t 3 others
 ORIGIN
 Query Match 21.0%; Score 372.6; DB 6; Length 646;
 Best Local Similarity 98.0%; Pred. No. 9.7e-75;
 Matches 398; Conservative 0; Mismatches 5; Indels 3; Gaps 2;
 QY 1369 ACAGGGAACAGATGAACCGCTCTGGAGAACACGTGCAGTCAACAAATGAAGTCTACA 1428
 Db 1 ACAGGGAACAGATGAACCGCTCTGGAGAACACGTGCAGTCAACAAATGAAGTCTACA 60
 QY 1429 GCTGGTAAACCATGAAGCAAAATCTGTGGCAGTGGATGTTATTCCTCATGTNGACAACCC 1488
 Db 61 GCTGGTAAACCATGAAGCAAAATCTGTGGCAGTGGATGTTATTCCTCATGTNGACAACCC 120
 QY 1489 AACCTTTGAAGAGACGAACTCCCTAATCAAGAGACTGCTGTTGAGAAATTAATCTTA 1548
 Db 121 AACCTTTGAAGAGACGAACTCCCTAATCAAGAGACTGCTGTTGAGAAATTAATCTTA 180
 QY 1549 AAATCTGTGTAATAGAAAATCTTGAACCATTTAGTAATAACAGAACTGCGCAATCAGGGCT 1608
 Db 181 AAATCTGTGTAATAGAAAATCTTGAACCATTTAGTAATAACAGAACTGCGCAATCAGGGCT 240
 QY 1609 AGTTTCTAATAAATTCGTAATAATTTAATAAATAGAGTGATGTAAGAGTCTCAG 1668
 Db 241 AGTTTCTAATAAATTCGTAATAATTTAATAAATAGAGTGATGTAAGAGTCTCAG 300
 QY 1669 ATGACTAATATATGCTATAGTTAAATTCGCTTAAATAATTTAACTGTTAACTTTT 1728
 Db 301 ATGACTAATATATGCTATAGTTAAATTCGCTTAAATAATTTAACTGTTAACTTTT 358
 QY 1729 CCACCAACTCATTAATAATTTTTCATAGGCAAGTTTCTCTCAG 1774
 Db 359 CCA-CAAACTCATTAATAATTTTTCATAGGCAAGTTTCTCTCAG 403
 RESULT 12
 AX396961/c

LOCUS
 DEFINITION Sequence 1176 from Patent WO0212328. linear PAT 18-MAY-2002
 ACCESSION AX396961
 VERSION AX396961.1 GI:21067708
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS King, G.E., Meagher, M.J., Xu, J., and Secrist, H.
 TITLE Compositions and methods for the therapy and diagnosis of colon cancer
 JOURNAL Patent: WO 0212328-A 1176 14-FEB-2002;
 CORIXA CORPORATION (US)
 FEATURES
 source Location/Qualifiers
 1..435
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 147 a 62 c 69 g 157 t
 ORIGIN
 Query Match 17.5%; Score 310; DB 6; Length 435;
 Best Local Similarity 97.4%; Pred. No. 2e-60;
 Matches 336; Conservative 0; Mismatches 6; Indels 3; Gaps 2;
 QY 1430 CTGTTAAACCATGAAGCAAAATTCGTGGCAGTGGATGTTATTCCTCATGTNGACAACCCA 1489
 Db 435 CTGTTAAACCATGAAGCAAAATTCGTGGCAGTGGATGTTATTCCTCATGTNGACAACCCA 376
 QY 1490 ACCTTTGAAGAGACGAACTCCCTAATCAAGAGACTGCTGTTGAGAAATTAATCTTAA 1549
 Db 375 ACCTTTGAAGAGACGAACTCCCTAATCAAGAGACTGCTGTTGAGAAATTAATCTTAA 316
 QY 1550 AATCTGTGTAATAGAAAATCTTGAACCATTTAGTAATAACAGAACTGCCATCAGGCTTA 1609
 Db 315 AATCTGTGTAATAGAAAATCTTGAACCATTTAGTAATAACAGAACTGCCATCAGGCTTA 256
 QY 1610 GTTCTTAAATAAATTTGGATAAAATTAATAAATGAAGTGTATCTGAAAGTGTCTCAGA 1669
 Db 255 GTTCTTAAATAAATTTGGATAAAATTAATAAATGAAGTGTATCTGAAAGTGTCTCAGA 196
 QY 1670 TGACTAATATATGCTATAGTTAAATGGCTTAAATAATTTAACTGTTAACTTTTTC 1729
 Db 195 TGACTAATATATGCTATAGTTAAATGGCTTAAATAATTTAACTGTTAACTTTTTC 138
 QY 1730 CACCAAACTCATTAATAATTTTTCATAGCAAGTTTCTCTCAG 1774
 Db 137 CA-CAAACTCATTAATAATTTTTCATAGCAAGTTTCTCTCAG 94
 RESULT 13
 AX115936/c
 LOCUS
 DEFINITION Sequence 195263 bp DNA linear HTG 20-JUN-2002
 ACCESSION AC115936
 VERSION AC115936.3 GI:21490425
 KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 195263)
 AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
 TITLE Mus musculus, clone RP24-545D3
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 195263)
 AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L.,
 Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
 Chazaro, B., Choepel, Y., Collangelo, M., Collins, S., Collamore, A.,
 Cook, A., Cooke, P., DeAtellano, K., Dewar, K., Diaz, J.S., Dodge, S.,


```

QY 494 GAGCCTGTGGTGGGCTCTGGTACCGCCGACGGGGCCCGGCGCTTAAACGCTCTTAAC 553
Db 179123 AAGAGATAGCAGGAGTGTGTGTGCCACACAGAGGG---AAAAATTCAAAATGCTTGTGAT 179067
QY 554 CCGCACACGAAATTCACGGTGCCACAGGTTTGGGGAGACCGTGGCAAGTCTCTTGGTTG 613
Db 179066 CCNATACCACTTTTATCTCTGCCACGG-----AACAAAGAGCCCTGGATC 179022
QY 614 GCCCTCATCAACCGCGCGGGGGTGCACCTTCGCACAGCAAGATCCATCTGGCTTATGAG 673
Db 179021 GCATCATTTGAAC---GGGAGGTTGTGCTTCACACAGAAAATTAAGGTGGCATCTGAG 178965
QY 674 AGATGGGGCTTGGAGCGCTCTGTTTAACTTCCCGGGACCGCGAATGAGTCACTCCCC 733
Db 178964 CACGAGCCACAGAGAGTGCATCTATATTTCCCAAGTACTCGCAACACGAGTGTTCCTCC 178905
QY 734 ATGTCTCACCGGGTGCAGTAGACATTTGTCATATGATCGGCAATCTGAAAGGCACA 793
Db 178904 ATGTCTCACAGGGTTTGAAGACATGTTGTAGTATGATTTGTAACATCAAGATGATG 178845
QY 794 AAAATCTGCAATCTTCAAGAGGCGATCAAGTGCACATGTCATAGAGTATGAGGAAA 853
Db 178844 GAGATTTTGCATTTAATTCGAAGGAGTTCATGTTACAGTCATGTTGGAGGTGGGAAGA 178785
QY 854 AAACATGCGCTTGGGTGAATCACTATTCAATTTTTTTCGTTCTGTGTCCTTTTATT 913
Db 178784 AAACATGTCATTTGGCTGAATCACTA-----TTTGTCTCTTTATGATG 178740
QY 914 ATTACGGCGGCACTGTGGGCTATTTTATCTTTTATCTGCTCGAAGGCTACGGAATGCA 973
Db 178739 GTCACACCGCCAGCTTAGCATATTTTCCCTTTTATCACATTCGGAGACTTTGGTGGCA 178680
QY 974 AGACCTCAAGCAGGAGCAGAGCAATTAAGCAGATGCTTAAAGACTATTTGAAAG 1033
Db 178679 AGAATTTGAGACAGGAGATGAAGCGATTAAACAGAGAGCTCAAGAAAGCTTTTGGCAG 178620
QY 1034 CTCAACTACGACACTGAACACAGGAGACAGGAATTCGCCCTGATGGAGATAGTTGT 1093
Db 178619 CTGCAAGTTCAGTATTAAGAAGGAGGATGAAGAGTAAATCAATGTCGACACTGT 178560
QY 1094 GCTGTGTGCAATGTAATTTATTAACCAAAATGATTTGGTAGCATCTTTAAAGTCAACCAT 1153
Db 178559 GTCATCTGCTTTGAAGCTTATAGCCTTAATGAATAGTTCGGATTTCTCACTGCAAACT 178500
QY 1154 ATTTTCCATAGACATGTTGTGACCATGCTGTGTTAAACACAGACTTCCCTCCATGTGC 1213
Db 178499 TTTTCCACAGAATTCGATTTGACCCCTGGATTTAGCCCAACGCGACATGCCCATGTGC 178440
QY 1214 AAATGTGACACTCAAGCTTTTGGGAATTTGAGGTGATGTTGAAGATGATCAGTGTCT 1273
Db 178439 AAATGTGACATTTGAAGCTTGGGATTCAGATGACATTTGAGGATGGACAGACTCT 178380
QY 1274 TTCAAGTCCCTGTATCAATGAATATCTAATAGTGCCTTCCCTCCCATGAAGAGGATA 1331
Db 178379 CTGCAAGTCTGATGTCAAATGAATTCGCCGAACCTTTGACCCGCGGAAGAAGAGA 178322

RESULT 14
AC111461/c
LOCUS AC111461 167021 bp DNA linear HTG 13-JUL-2002
DEFINITION Rattus norvegicus clone CH230-213C21, *** SEQUENCING IN PROGRESS
***, 75 unordered pieces.
ACCESSION AC111461
VERSION AC111461.2 GI:21736011
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 167021)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,

```

TITLE JOURNAL REFERENCE AUTHORS TITLE JOURNAL REFERENCE AUTHORS TITLE JOURNAL COMMENT

Albrooke, S.L., Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T.,
Barbaria, J., Benton, J., Bimge, K., Blankenburg, K., Bonnin, D.,
Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
Carroll, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
Devila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, N.C.,
Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
Gabibi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
Gorell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,
Honsi, F., Howard, S., Huber, J., Huly, S., Hume, J., Jackson, L.E.,
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
Karleson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loubege, H.,
Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okwuonu, G.,
Ogundimu, N., Oviado, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G.,
Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sleson, I.,
Sodergren, E., Sonaika, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Umani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wlezyk, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G., and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 167021)
Worley, K.C.
Direct Submission
Submitted (19-FEB-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 167021)
Worley, K.C.
Direct Submission
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 12, 2002 this sequence version replaced gi:18701226.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GMJK
Center clone name: CH230-213C21
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 89039 bases at least Q40
Consensus quality: 93636 bases at least Q30
Consensus quality: 97431 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently

- * consists of 75 contigs. The true order of the pieces
- * is not known and their order in this sequence record is
- * arbitrary. Gaps between the contigs are represented as
- * runs of N, but the exact sizes of the gaps are unknown.
- * This record will be updated with the finished sequence
- * as soon as it is available and the accession number will
- * be preserved.

1	1200: contig of 1200 bp in length
1201	1300: gap of unknown length
1301	2480: contig of 1180 bp in length
2481	2580: gap of unknown length
2581	4294: contig of 1714 bp in length
4295	4394: gap of unknown length
4395	5589: contig of 1195 bp in length
5590	5689: gap of unknown length
5690	7039: contig of 1350 bp in length
7040	7139: gap of unknown length
7140	8161: contig of 1022 bp in length
8162	8261: gap of unknown length
8262	9786: contig of 1525 bp in length
9787	9886: gap of unknown length
9887	11075: contig of 1189 bp in length
11076	11175: gap of unknown length
11176	12487: contig of 1312 bp in length
12488	12587: gap of unknown length
12588	13739: contig of 1152 bp in length
13740	13839: gap of unknown length
13840	15104: contig of 1265 bp in length
15105	15204: gap of unknown length
15205	16606: contig of 1402 bp in length
16607	16706: gap of unknown length
16708	18097: contig of 1391 bp in length
18098	18197: gap of unknown length
18198	19766: contig of 1569 bp in length
19767	19866: gap of unknown length
19867	21733: contig of 1869 bp in length
21736	21835: gap of unknown length
21836	22871: contig of 1036 bp in length
22872	22971: gap of unknown length
22972	24060: contig of 1089 bp in length
24061	24160: gap of unknown length
24161	25579: contig of 1419 bp in length
25580	25679: gap of unknown length
25680	27146: contig of 1467 bp in length
27147	27246: gap of unknown length
27247	28388: contig of 1142 bp in length
28389	28488: gap of unknown length
28489	30073: contig of 1585 bp in length
30074	30173: gap of unknown length
30174	31407: contig of 1234 bp in length
31408	31507: gap of unknown length
31508	33049: contig of 1542 bp in length
33050	33149: gap of unknown length
33150	33485: contig of 1707 bp in length
33485	34956: gap of unknown length
34957	36203: contig of 1247 bp in length
36204	36303: gap of unknown length
36304	38026: contig of 1723 bp in length
38027	38126: gap of unknown length
38127	39131: contig of 1005 bp in length
39132	39231: gap of unknown length
39232	40881: contig of 1650 bp in length
40882	40981: gap of unknown length
40982	42222: contig of 1241 bp in length
42223	42322: gap of unknown length
42323	43979: contig of 1657 bp in length
43980	44079: gap of unknown length
44080	45706: contig of 1627 bp in length
45707	45806: gap of unknown length
45807	47463: contig of 1657 bp in length
47464	47563: gap of unknown length
47564	49070: contig of 1507 bp in length
49071	49170: gap of unknown length

Query Match 15.7%; Score 278.4; DB 2; Length 167021;
Best Local Similarity 60.7%; Pred. No. 1.3e-52;
Matches 54; Conservative 0; Mismatches 316; Indels 36;

Qy	434	GTGAACCGTAGCGGTGTGGAGCTGAGCGAGAGAGGGGTGTACGCCACGAGACTCGCGCTG	493
Db	83860	GGGAACCGCATCTGTACAGCTGGCGAGACTGGGGTCTTCGGAAGAAGACTCCATCTTG	83801
Qy	494	GAGCCTGTGGCTGGGGTCTCGGTACGCCCGA CGGGCCCGGGCGCTTAACGCTGTAAAC	553
Db	83800	AAGAGAGTAGCAGGAGTAGTTGTGCCACACAGAGGAAA --- AATTCAAAATGCTTTGTGAC	83744
Qy	554	CCGCACACGAATTTTCA CGGTGCCACCGTTTGGGGAAGCACCGTGCAGTCTCTCTTGGTTG	613
Db	83743	CCCAATACCAGTTTCATCTGCGCGGAA -----CAAAGAGCGGTGGATT	83699
Qy	614	GCCTCATCAACGGCGGGGGCTGCCTTGCAGACAAGATCCATCTGGCTTTATGAG	673
Db	83698	GCACCTCAITGAA ---AGGGAGAGTTGTGCGCTTCACACAGAAAATCAAAGTGGCATCTGAG	83642
Qy	674	AGATGGGCGTCTGGAGCGCTCATCTTTAACTTCCCGGAGCCGGAATGAGGTCAATCCCC	733
Db	83641	AATGGGGCCAGGGAGTGATCATCTATACTTTCCGGGGACTGGCAACACAGGTTTTTCCCC	83582
Qy	734	ATGTCTCAACCGGGTGCAGTAGACAATTTGTTGCAATCATGATCGGCAATCTGAAGAGGCACA	793
Db	83581	ATGTCTCAACGAGCGTTTGAAGACATCGTTGTAGTTATGTTGGTAAACGTCAAAGGTATG	83522
Qy	794	AAAAATCTGCAATCTATTCAAGAGGCGATACAAGTGACAAATGGTGCATAGAAGTAGGGAAA	853
Db	83521	GAGATTTTGCAATTAATTAGGAAGGAGTCCATGTTACGGTTCATGTTGAGGTGGGAAGA	83462

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 26, 2003, 00:41:20 ; Search time 284 Seconds
(without alignments)
14067.060 Million cell updates/sec

Title: US-09-854-300-7

Perfect score: 1774

Sequence: 1 agctggagctccacggcgt.....ataggcaagttctctctcag 1774

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 101002:*

- 1: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1773	99.9	1774	AAI72095	Human GRAIL cDNA.
2	1662.8	93.7	2773	AAAI5995	Human protein clon
3	1662.8	93.7	2796	AAI72321	ISIGP-4 cDNA. Hom
4	1431.4	80.7	2876	AAAS42511	Human cDNA encodin
5	1278.2	72.1	1284	AAAI5985	Human protein clon
6	1178.8	66.4	2306	AAAS25884	Human cDNA encodin
7	1167.2	65.8	1249	ABL90796	Human polynucleoti
8	1166	65.7	2145	AAI72094	Mouse GRAIL cDNA.
9	1157.6	65.3	1250	AAAS26340	Human cDNA encodin

	10	1106.8	62.4	1287	24	ABK12983
	11	1052.4	59.3	1287	24	ABK12992
	12	1042.8	58.8	1287	24	ABK12990
	13	1042.8	58.8	1287	24	ABK12991
	14	1019.4	57.5	1135	22	AAI59292
	15	998.6	56.3	1287	24	ABK12989
	16	977.2	55.1	1287	24	ABK12987
	17	972.4	54.8	1287	24	ABK12988
	18	655.2	36.9	737	22	AAI61078
	19	497.2	28.0	502	24	ABN19991
	20	392.2	22.1	404	22	AAF66696
	21	372.6	21.0	646	24	ABL38226
	22	310	17.5	435	24	ABK45625
	23	254.8	14.4	1355	21	AAAB7765
	24	254.8	14.4	1355	22	AAF64047
	25	243.8	13.7	918	23	AAI70124
	26	210.2	11.8	2826	24	ABL90797
	27	209	11.8	1200	22	AAF94484
	28	209	11.8	1527	22	AAF94494
	29	209	11.8	2005	22	AAK94499
	30	209	11.8	2063	19	AAV43601
	31	209	11.8	2452	22	AAK94373
	32	194.8	11.0	1869	22	ABA09101
	33	194.8	11.0	1869	22	AAK51507
	34	194.8	11.0	1869	22	AAK52491
	35	191.8	10.8	1708	23	AAK76863
	36	180.8	10.2	3615	24	ABK36205
	37	178.8	10.1	631	21	AAK76002
	38	149.2	8.4	1322	19	AAV43616
	39	148.8	8.4	1260	24	ABK12986
	40	147.6	8.3	1253	19	AAV40387
	41	147.6	8.3	1257	22	AAF94418
	42	147.6	8.3	1433	22	AAF94428
	43	147.6	8.3	1489	24	ABL55072
	44	147.6	8.3	1515	24	ABL90798
	45	147.6	8.3	1546	24	ABL55094

ALIGNMENTS

RESULT 1

AAI72095
ID AAI72095 standard; cDNA; 1774 BP.

AC AAI72095;

DT 25-MAR-2002 (first entry)

DE Human GRAIL cDNA.

XX Murine; human; GRAIL; anergy; attenuation; tyrosine; phosphorylation;
XX antigenic stimulation; interleukin-2; gene therapy; polymorphism; IL-2;
XX autoimmune disease; tumour cell; cancer; transplant rejection; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 263..1547

FT /*tag= a

FT /product= "GRAIL"

XX WO200185943-A1.

XX PD 15-NOV-2001.

XX PF 11-MAY-2001; 2001WO-US15385.

XX PR 11-MAY-2000; 2000US-203513P.

XX (STRD) UNIV LELAND STANFORD JUNIOR.

XX Ford GS, Bloom D, Fathman CG;

PI

Db 1441 TGAAGCAAAATCTGTCGAGTGGATGTTATTTCTCATGTGNGACAACCCAACTTTGAAGA 1500
 QY 1501 AGACGAAACTCTTAATCAAGAGACTGCTGTTCGAGAAATTAATCTTAAATCTGTGTAA 1560
 Db 1501 AGACGAAACTCTTAATCAAGAGACTGCTGTTCGAGAAATTAATCTTAAATCTGTGTAA 1560
 QY 1561 ATAGAAACTTGACCAATAGTAATAACAGAACTGCCAATCAGGGCTAGTTTCTATTAA 1620
 Db 1561 ATAGAAACTTGACCAATAGTAATAACAGAACTGCCAATCAGGGCTAGTTTCTATTAA 1620
 QY 1621 TAAATTCGATAAAATTAATAAATAAGAGTGATACTGAAAGTGTCTCAGATGACTATATTT 1680
 Db 1621 TAAATTCGATAAAATTAATAAATAAGAGTGATACTGAAAGTGTCTCAGATGACTATATTT 1680
 QY 1681 ATGCTATAGTAAATGGCTAAATAATTAACCTGTTAACTTTTTCACCAAACTCA 1740
 Db 1681 ATGCTATAGTAAATGGCTAAATAATTAACCTGTTAACTTTTTCACCAAACTCA 1740
 QY 1741 TTATAATATTTTTCATAGGCAAGTTTCTCTCAG 1774
 Db 1741 TTATAATATTTTTCATAGGCAAGTTTCTCTCAG 1774

RESULT 2
 ID AAA15995 standard; cDNA; 2773 BP.
 XX
 AC AAA15995;
 XX
 DT 12-JUN-2000 (first entry)
 XX
 DE Human protein clone HP10574 full length coding sequence.
 XX
 KW Human protein; hydrophobic domain; nutritional source; haematopoiesis;
 KW cytokine production; cell proliferation; cell differentiation;
 KW immune deficiency; infectious disease; autoimmune disorder; asthma;
 KW multiple sclerosis; systemic lupus erythematosus; rheumatoid arthritis;
 KW allergic reaction; osteoporosis; osteoarthritis; periodontal disease;
 KW nervous system disorder; Alzheimer's disease; Parkinson's disease;
 KW Huntington's disease; liver fibrosis; lung fibrosis; reperfusion injury;
 KW systemic cytokine damage; tissue differentiation; contraceptive; stroke;
 KW coagulation disorder; myocardial infarction; inflammatory condition;
 KW septic shock; sepsis; ischaemia; reperfusion injury; arthritis; tumour;
 KW nephritis; therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200005367-A2.
 XX
 PD 03-FEB-2000.
 XX
 PF 22-JUL-1999; 99WO-JF03929.
 XX
 PR 24-JUL-1998; 98JP-0208820.
 PR 07-AUG-1998; 98JP-0224105.
 PR 25-AUG-1998; 98JP-0238116.
 PR 09-SEP-1998; 98JP-0254736.
 PR 29-SEP-1998; 98JP-0275505.
 XX
 PA (SAGA) SAGAMI CHEM RES CENT.
 PA (PROT-) PROTEGENE INC.
 XX
 PI Kato S, Kimura T;
 XX
 DR WPI; 2000-182694/16.
 DR P-PSDB; AAY94897.
 XX
 PT Novel human proteins having hydrophobic domains useful for treating
 PT osteoporosis, Alzheimer's disease, Parkinson's disease, asthma,
 PT multiple sclerosis, rheumatoid arthritis, cancer, anaemia, and stroke -
 XX
 PS Claim 4; Page 348-351; 351pp; English.
 XX

CC This sequence encodes a human protein of the invention, which has
 CC hydrophobic domains. The DNA sequences can be used as a probe or as a
 CC genetic marker. The protein can also be used as a marker, and to identify
 CC potential genetic disorders. The DNA and protein can also be used as
 CC nutritional sources or supplements. The protein exhibits cytokine, cell
 CC proliferation, cell differentiation activities and induces production of
 CC other cytokines in certain cell populations. The protein also exhibits
 CC immunostimulating or immune suppressing activity. It can be used in the
 CC treatment of various immune deficiencies and disorders, and to treat
 CC infectious diseases caused by viral, bacterial, fungal or other
 CC infections. The protein is also used for treating autoimmune disorders
 CC such as multiple sclerosis, systemic lupus erythematosus, and rheumatoid
 CC arthritis. It is also useful in the treatment of allergic reactions and
 CC conditions such as asthma, and in immune suppression after organ
 CC transplantation. The protein is useful in regulation of haematopoiesis
 CC and consequently in the treatment of myeloid or lymphoid cell
 CC deficiencies. It is also used in compositions for tissue growth or
 CC regeneration. The protein is also used in the treatment of osteoporosis
 CC or osteoarthritis and in the treatment of periodontal disease and other
 CC tooth repair processes. The protein is used in the treatment of nervous
 CC system disorders such as Alzheimer's disease, Parkinson's disease, and
 CC Huntington's disease. They are useful for protection or regeneration and
 CC treatment of lung or liver fibrosis, reperfusion injury in various
 CC tissues, and conditions resulting from systemic cytokine damage. They are
 CC also used for promoting or inhibiting tissue differentiation. They are
 CC also used as contraceptives since they exhibit activin or inhibin related
 CC activities and as a fertility inducing therapeutic. They are used for
 CC treating various coagulation disorders and in treatment and prevention of
 CC conditions resulting from coagulation activities e.g. myocardial
 CC infarction or stroke. They also acts as receptors, receptor ligands or
 CC inhibitors or agonists of receptor/ligand interactions. They are used to
 CC treat inflammatory conditions such as septic shock, sepsis, ischaemia
 CC reperfusion injury, arthritis, and nephritis. They can be used to
 CC prevent tumours.
 XX

SQ Sequence 2773 BP; 803 A; 545 C; 617 G; 808 T; 0 other;

Query Match 93.7%; Score 1662.8; DB 21; Length 2773;

Best Local Similarity 99.4%; Pred. No. 0;

Matches 1690; Conservative 0; Mismatches 8; Indels 3; Gaps 2;

QY 74 CGAGAGAGCTGATCTGGCGCAACTGTGCTGCTGACGCTAGCTCTCTCGCTCCGACG 133

Db 22 CGAGAGAGCTGATCTGGCGCAACTGTGCTGCTGACGCTAGCTCTCTCGCTCCGACG 81

QY 134 TAGCTCGAGCTCCCGAGTCTCACTCCATTCTTCCACCTCGCGGCGACCTGCTCAAG 193

Db 82 TAGCTCGAGCTCCCGAGTCTCACTCCATTCTTCCACCTCGCGGCGACCTGCTCAAG 141

QY 194 ACCAGGCTCTGCGCAAGCGCTAGGAGGCGGCTGCGAGGGGCGCTAGGGAACCTGGGAGC 253

Db 142 ACCAGGCTCTGCGCAAGCGCTAGGAGGCGGCTGCGAGGGGCGCTAGGGAACCTGGGAGC 201

QY 254 GCGCGCGCATGGGGGCGCGCGCTGGGCGCGGGGTCTCTCTCCCGCGGTGGCTGGCGGCTTT 313

Db 202 GCGCGCGCATGGGGGCGCGCGCTGGGCGCGGGGTCTCTCTCCCGCGGTGGCTGGCGGCTTT 261

QY 314 TCCAGATTGCTGGCATGTGTCTTCTGTGCTGCTTCCAGAGGACACCGGTTCCCGG 373

Db 262 TCCAGATTGCTGGCATGTGTCTTCTGTGCTGCTTCCAGAGGACACCGGTTCCCGG 321

QY 374 GGGGCTGAAGCAGTGTGGACCGCGTACCTCAACGTGTCTCGCGGGTTCGCGACACGGGA 433

Db 322 GGGGCTGAAGCAGTGTGGACCGCGTACCTCAACGTGTCTCGCGGGTTCGCGACACGGGA 381

QY 434 GTGAACCGTACGCTGGGAGCTGAGCGAGGAGGCGGTGTGTACGCGCAGGACTCGCGCGTG 493

Db 382 GTGAACCGTACGCTGGGAGCTGAGCGAGGAGGCGGTGTGTACGCGCAGGACTCGCGCGTG 441

QY 494 GAGCCTGTGGTGGGCTCTGTGTACCGCCCGACCGGGCCCGGGCGCTTTAAACGCTGTAAAC 553

Db 442 GAGCCTGTGGTGGGCTCTGTGTACCGCCCGACCGGGCCCGGGCGCTTTAAACGCTGTAAAC 501

QY 554 CCGCACAGAAATTTTCAACGGTGGCCCAACGGTTTGGGGAAGACACCGTGCAGTCTCTTGGTTG 613
DQ 502 CCGCACAGAAATTTTCAACGGTGGCCCAACGGTTTGGGGAAGACACCGTGCAGTCTCTTGGTTG 561
QY 614 GGCCTCATCAACCGGGGGGGGCTGCACCTTCCGAGACAAGATCCATCTGGCTTATGAG 673
DQ 562 GGCCTCATCAACCGGGGGGGGCTGCACCTTCCGAGACAAGATCCATCTGGCTTATGAG 621
QY 674 AGATGGGGCTCTGGAGCGCTCATCTTTAACTTCCCGGGGACCCGCAATGAGTGCATCCCC 733
DQ 622 AGAGGGGGCTCTGGAGCGCTCATCTTTAACTTCCCGGGGACCCGCAATGAGTGCATCCCC 681
QY 734 ATGTCTACCCGGGTGCAGTATGATGATTTGCAATCATGATCGGCAATCTGAAAGGCACA 793
DQ 682 ATGTCTACCCGGGTGCAGTATGATGATTTGCAATCATGATCGGCAATCTGAAAGGCACA 741
QY 794 AAAATTTCTGCAATCTTAAAGAGGCATACAGTGCATGATGATGATGATGATGATGATGAT 853
DQ 742 AAAATTTCTGCAATCTTAAAGAGGCATACAGTGCATGATGATGATGATGATGATGATGAT 801
QY 854 AAAATTTCTGCAATCTTAAAGAGGCATACAGTGCATGATGATGATGATGATGATGATGAT 913
DQ 802 AAAATTTCTGCAATCTTAAAGAGGCATACAGTGCATGATGATGATGATGATGATGATGAT 861
QY 914 ATTACGGGGCAACCTGTGGGCTATTTTATCTTTTATCTGCTCGAAGGCTACGGAATGCA 973
DQ 862 ATTACGGGGCAACCTGTGGGCTATTTTATCTTTTATCTGCTCGAAGGCTACGGAATGCA 921
QY 974 AGAGTCAAAGCAGGAGCAGAGGCAATTAAGGCGAGATGCTAAAAAGCTATTGGAAGG 1033
DQ 922 AGAGTCAAAGCAGGAGCAGAGGCAATTAAGGCGAGATGCTAAAAAGCTATTGGAAGG 981
QY 1034 CTTCAACTACGACACTGAAACAGAGGAGCAAGGAAATTTGGCCCTGATGAGATAGTGT 1093
DQ 982 CTTCAACTACGACACTGAAACAGAGGAGCAAGGAAATTTGGCCCTGATGAGATAGTGT 1041
QY 1094 GCTGTGTGCTTGAATCTTAAACCAATGATTTGGTACGATCTTAAAGTGCACCAAT 1153
DQ 1042 GCTGTGTGCTTGAATCTTAAACCAATGATTTGGTACGATCTTAAAGTGCACCAAT 1101
QY 1154 ATTTTCCATAAGACATGTTGTGACCCATGGCTGTTTAAACCAAGACTTTGCCCATGTGC 1213
DQ 1102 ATTTTCCATAAGACATGTTGTGACCCATGGCTGTTTAAACCAAGACTTTGCCCATGTGC 1161
QY 1214 AAATGTGACATCACTCAAGCTTTGGGAATTTGAGTGGATGTTGAGATGATGATGATGAT 1273
DQ 1162 AAATGTGACATCACTCAAGCTTTGGGAATTTGAGTGGATGTTGAGATGATGATGATGAT 1221
QY 1274 TTACAAGTCCCTGTATCCAAATGAAATATCTAATAGTGCCTCTCCCATGAAGAGATAAT 1333
DQ 1222 TTACAAGTCCCTGTATCCAAATGAAATATCTAATAGTGCCTCTCCCATGAAGAGATAAT 1281
QY 1334 CCGAGGAGACCGCATCATCTGGATATGCTTACGATGAGGAAACAGATGAAACCGCTCTG 1393
DQ 1282 CCGAGGAGACCGCATCATCTGGATATGCTTACGATGAGGAAACAGATGAAACCGCTCTG 1341
QY 1394 GAGGAACAGCTGCAGTCAACAAATGAAATGCTTACAGCTGGTAAACCATGAAGCAATCT 1453
DQ 1342 GAGGAACAGCTGCAGTCAACAAATGAAATGCTTACAGCTGGTAAACCATGAAGCAATCT 1401
QY 1454 GTGGCAGTGGATGTTTATCTCTCATGTNGACACACCAACCTTTGAGAGAGAGCAAACTCT 1513
DQ 1402 GTGGCAGTGGATGTTTATCTCTCATGTNGACACACCAACCTTTGAGAGAGAGCAAACTCT 1461
QY 1514 AATCAAGACATCTGCTGTCGAGAAATTAATCTTAAATCTGTGTAAATAGAAACTTGA 1573
DQ 1462 AATCAAGACATCTGCTGTCGAGAAATTAATCTTAAATCTGTGTAAATAGAAACTTGA 1521
QY 1574 ACCATTAGTAATAACAGAACTGCCAATCAGGGCCCTAGTCTTATTAATAAATTTGATAAA 1633
DQ 1522 ACCATTAGTAATAACAGAACTGCCAATCAGGGCCCTAGTCTTATTAATAAATTTGATAAA 1581
QY 1634 TTTAATAAATAAGAGTGATGATGCTGAAAGTGCCTCAGATGACTAATATTATGCTATAGTTAA 1693

DQ 1582 TTTAATAAATAAGAGTGATGATGCTGAAAGTGCCTCAGATGACTAATATTATGCTATAGTTAA 1641
QY 1694 AATGCTTTAAAAATATTTAACTGTTAACTTTTTCACCAAACTCATTATAATATTTT 1753
DQ 1642 A--TGCTTAAAAATATTTAACTGTTAACTTTTTCCTCA--CAAACTCATTATAATATTTT 1698
QY 1754 CATAGGCAAGTTTCTCTCTCAG 1774
DQ 1699 CATAGGCAAGTTTCTCTCTCAG 1719
RESULT 3
AAI72321
ID AAI72321 standard; cDNA; 2796 BP.
XX AAI72321;
AC AAI72321;
XX 15-APR-2002 (first entry)
XX ISIGP-4 cDNA.
XX Human; intracellular signalling protein; ISIGP; gene; cell proliferation;
KW autoimmune; inflammation; gastrointestinal disorder;
KW reproductive disorder; developmental disorder; ss.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 217..1503
FT /*tag= a
FT /product= "ISIGP-4"
XX WO200194391-A2.
XX 13-DEC-2001.
XX 07-JUN-2001; 2001WO-US18595.
XX 08-JUN-2000; 2000US-210582P.
XX 16-JUN-2000; 2000US-212443P.
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Yue H, He A, Nguyen DB, Yao MG, Bandman O, Burford N, Tang YT;
PI Xu Y, Hatalia A, Azimzai Y, Wallia NK;
XX
XX WPI; 2002-154564/20.
DR P-PSDB; AAB47874.
XX
XX New human intracellular signaling protein and polynucleotides useful
PT for diagnosing, treating or preventing cell proliferative,
PT autoimmune/inflammatory, gastrointestinal, reproductive and
PT developmental disorders -
XX
PS Claim 5; Page 104; 106pp; English.
XX
XX The sequences given in AAI72318-22 encode novel human intracellular
CC signalling proteins (ISIGP). The polynucleotides and ISIGP proteins
CC may be used for the diagnosis, treatment or prevention of cell
CC proliferative, autoimmune/inflammatory, gastrointestinal, reproductive
CC and developmental disorders. The protein encoded by this sequence has
CC homology to mouse gl-related zinc finger protein.
XX
SQ Sequence 2796 BP; 821 A; 548 C; 620 G; 807 T; 0 other;
Query Match 93.7%; Score 1662.8; DB 24; Length 2796;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1690; Conservative 0; Mismatches 8; Indels 3; Gaps 2;
QY 74 CCGAGAGCTGCATCTGCGGCAACCTGTGTGCTGACGCTACGCTCTCTGCTCGGAG 133
DQ 28 CCGAGAGCTGCATCTGCGGCAACCTGTGTGCTGACGCTACGCTCTCTGCTCGGAG 87

```

QY 134 TAGCTGCGAGCTCCCGAGTCTCACTCCATTCCTTCCCACTCTGGCGCGCACTGCTCAAG 193
DB 88 TAGCTGCGAGCTCCCGAGTCTCACTCCATTCCTTCCCACTCTGGCGCGCACTGCTCAAG 147
QY 194 ACCAGGGTCTCTGCAAGCGCTAGGAGGGCGGTGCGAGGGCGCTAGGGAACATGCGGAGC 253
DB 148 ACCAGGGTCTCTGCAAGCGCTAGGAGGGCGGTGCGAGGGCGCTAGGGAACATGCGGAGC 207
QY 254 GCGCGGCCATGAGGGCGCGGCTGCGGCGCGGGTCTCTCGCGGGGTGGCTGCGGCTTT 313
DB 208 GCGCGGCCATGAGGGCGCGGCTGCGGCGCGGGTCTCTCGCGGGGTGGCTGCGGCTTT 267
QY 314 TCCAGATTGCTGSCATGCTGCTTCTGCTGCGCCCTGAGTTCGAGACAGCCCGGTTCGCGG 373
DB 268 TCCAGATTGCTGSCATGCTGCTTCTGCTGCGCCCTGAGTTCGAGACAGCCCGGTTCGCGG 327
QY 374 GGGGCTGAAGAGCAGTGTGGACCGGCTACCTCAACGCTGCTGCGGGTTCGCGCACAGGGA 433
DB 328 GGGGCTGAAGAGCAGTGTGGACCGGCTACCTCAACGCTGCTGCGGGTTCGCGCACAGGGA 387
QY 434 GTGAACCGTACGCTGTGGAGCTGAGCGAGGAGGGCGGTGACGCGCAGGACTCGCGCTG 493
DB 388 GTGAACCGTACGCTGTGGAGCTGAGCGAGGAGGGCGGTGACGCGCAGGACTCGCGCTG 447
QY 494 GAGCCTGTGGCTGGGCTGCTGCTACCGCCGACGCGGCCGCGGGCGCTTAAACCTGTAAAC 553
DB 448 GAGCCTGTGGCTGGGCTGCTGCTACCGCCGACGCGGCCGCGGGCGCTTAAACCTGTAAAC 507
QY 554 CCGCACACGAAATTCACGGTGCACCGTTCGCGGAAGCACCGTGCAGCTCTCTTGGTTG 613
DB 508 CCGCACACGAAATTCACGGTGCACCGTTCGCGGAAGCACCGTGCAGCTCTCTTGGTTG 567
QY 614 GCCTCATCAACGCGCGGGGCTGCTGACCTTTCGCGACAGAAATCCATCGCTTATGAG 673
DB 568 GCCTCATCAACGCGCGGGGCTGCTGACCTTTCGCGACAGAAATCCATCGCTTATGAG 627
QY 674 AGATGGCGCTCTGGAGCGCTCATCTTTAACTTCCCGGACCGCAATGAGTGCATCCCC 733
DB 628 AGAGGGCGCTCTGGAGCGCTCATCTTTAACTTCCCGGACCGCAATGAGTGCATCCCC 687
QY 734 ATGCTCACCGGGTGCAGTAGACATTTGTGCAATCATGATCGGCAATCTGAAAGGCACA 793
DB 688 ATGCTCACCGGGTGCAGTAGACATTTGTGCAATCATGATCGGCAATCTGAAAGGCACA 747
QY 794 AAAATTCTGCAATCTATTCAAGAGGCATACAAAGTGACAAATGGTGCATAGAGTAGGAAA 853
DB 748 AAAATTCTGCAATCTATTCAAGAGGCATACAAAGTGACAAATGGTGCATAGAGTAGGAAA 807
QY 854 AAACATGGCCCTTGGGTGAATCACTATTCAATTTTTTTTCTGTTCTGTCCTTTTATT 913
DB 808 AAACATGGCCCTTGGGTGAATCACTATTCAATTTTTTTTCTGTTCTGTCCTTTTATT 867
QY 914 ATTACGGCGCAACTGTGGCTATTTTATCTTTTATCTGCTCGAAGGCTACGGAATGCA 973
DB 868 ATTACGGCGCAACTGTGGCTATTTTATCTTTTATCTGCTCGAAGGCTACGGAATGCA 927
QY 974 AGAGCTCAAGCAGGAGCAGAGGCAATTAAGGCGAGTGTCTAAAAAGCTATTGGAAG 1033
DB 928 AGAGCTCAAGCAGGAGCAGAGGCAATTAAGGCGAGTGTCTAAAAAGCTATTGGAAG 987
QY 1034 CTTCAACTACGCACACTGAAACAAAGGAGCAAGGAATTTGGCCTGATGGAGATAGTTGT 1093
DB 988 CTTCAACTACGCACACTGAAACAAAGGAGCAAGGAATTTGGCCTGATGGAGATAGTTGT 1047
QY 1094 GCTGTGTGCAATTGAAATTGATATAACCAAAATGATTTGGTACGCACTTTAAACGTCGAACCAT 1153
DB 1048 GCTGTGTGCAATTGAAATTGATATAACCAAAATGATTTGGTACGCACTTTAAACGTCGAACCAT 1107
QY 1154 ATTTTCCATAAGACATGCTGTGACCCATGCTGCTTTAAACACAGACTTTGCCCATGTGC 1213
DB 1108 ATTTTCCATAAGACATGCTGTGACCCATGCTGCTTTAAACACAGAGACTTTGCCCATGTGC 1167

```

```

QY 1214 AAATGTGACATATCAAAAGCTTTGGAAATTTGAGGTGGATGTTGAAGATGGATCAGTGTCT 1273
DB 1168 AAATGTGACATATCAAAAGCTTTGGAAATTTGAGGTGGATGTTGAAGATGGATCAGTGTCT 1227
QY 1274 TTACAGTCTCCTGTATCCAAATGAATATCTAATAGTGCCTCTCCCATGAGAGGATAAT 1333
DB 1228 TTACAGTCTCCTGTATCCAAATGAATATCTAATAGTGCCTCTCCCATGAGAGGATAAT 1287
QY 1334 CGCAGCGAGACCGGCATCATCTGGATATGCTTCAAGTACAGGGAACAGATGAACCGCCTCTG 1393
DB 1288 CGCAGCGAGACCGGCATCATCTGGATATGCTTCAAGTACAGGGAACAGATGAACCGCCTCTG 1347
QY 1394 GAGGAACACGCTGACGATCAACAAATGAAAGTCTACAGCTGGTAAACCATGAAGCAAAATCT 1453
DB 1348 GAGGAACACGCTGACGATCAACAAATGAAAGTCTACAGCTGGTAAACCATGAAGCAAAATCT 1407
QY 1454 GTGCGAGTGGATGTTATTCCTCATGTGACCAACCCCACTTTTCAAGAGACGAACTCCT 1513
DB 1408 GTGCGAGTGGATGTTATTCCTCATGTGACCAACCCCACTTTTCAAGAGACGAACTCCT 1467
QY 1514 AATCAAGAGACTGCTGTTTCGAGAAATTTAAATCTTTAAATCTGTGTTAAATAGAAAACTTGA 1573
DB 1468 AATCAAGAGACTGCTGTTTCGAGAAATTTAAATCTTTAAATCTGTGTTAAATAGAAAACTTGA 1527
QY 1574 ACCATTAGTAAATACAGAACTGCAATCAGGGCTAGTTTCTATTATTAATTAATTTGGATAAA 1633
DB 1528 ACCATTAGTAAATACAGAACTGCAATCAGGGCTAGTTTCTATTATTAATTAATTTGGATAAA 1587
QY 1634 TTTAATAAATAAGAGTGACTGAAAGTCTCAGATGACTAATATTATGCTATAGTTAA 1693
DB 1588 TTTAATAAATAAGAGTGACTGAAAGTCTCAGATGACTAATATTATGCTATAGTTAA 1647
QY 1694 AATGGCTTAAAAATATTTAACTGTTTAACTTTTCCACCAAACTCATTTATAATATTTT 1753
DB 1648 A--TGGCTTAAAAATTTTAACTGTTTAACTTTTTCCTCA--CAAACTCATTATAATATTTT 1704
QY 1754 CATAGGCAAGTTTCTCTCAG 1774
DB 1705 CATAGGCAAGTTTCTCTCAG 1725

```

RESULT 4

AAS42511

ID AAS42511 standard; cDNA; 2876 BP.

XX AAS42511;

XX 18-DEC-2001 (first entry)

XX Human cDNA encoding an mddt protein, clone LI:757439.1:2000MAY01.

Human; molecules for disease detection and treatment; mddt; ss;
 Antiarteriosclerotic; hepatotropic; antipsoriatic; antiproliferative;
 immunosuppressive; antidiabetic; antiaspartic; neuroprotective;
 osteopathic; antihypertensive; cell proliferative disorder;
 arteriosclerosis; cirrhosis; psoriasis; cancer; adenocarcinoma;
 leukemia; breast cancer; autoimmune disorder; AIDS;
 acquired immunodeficiency syndrome; Addison's disease;
 diabetes mellitus; asthma; multiple sclerosis; osteoarthritis.

XX Homo sapiens.

OS XX

PN XX

XX W0200162922-A2.

XX 30-AUG-2001.

XX 21-FEB-2001; 2001WO-US05896.

XX 24-FEB-2000; 2000US-0185213.

XX 16-MAY-2000; 2000US-0205232.

XX 17-MAY-2000; 2000US-0205285.

XX 17-MAY-2000; 2000US-0205286.

XX 17-MAY-2000; 2000US-0205287.

Db 661 GCAACTGTGGGCTATTTATCTTTTATCTCTGCTCAAGGCTACGGAATGCAAGAGCTCAA 720
Qy 983 AGCAGGAAGCAGAGGCAATTAAGGCAGATGCTAAAAAGCTATTGGAAGCTTCAACTA 1042
Db 721 AGCAGGAAGCAGAGGCAATTAAGGCAGATGCTAAAAAGCTATTGGAAGCTTCAACTA 780
Qy 1043 CGCACACTGAACAAGGAGACAAGGAAATTTGGCCCTGATGAGATAGTTGTGCTGTGTC 1102
Db 781 CGCACACTGAACAAGGAGACAAGGAAATTTGGCCCTGATGAGATAGTTGTGCTGTGTC 840
Qy 1103 ATTGAATTGTATAAACCAATGATTTGTTAGTCATCTTAAAGTGCACCATATTTTCCAT 1162
Db 841 ATTGAATTGTATAAACCAATGATTTGTTAGTCATCTTAAAGTGCACCATATTTTCCAT 900
Qy 1163 AAGACATGTTGTGACCCCATGCTGTTAAACACAAAGACTTGGCCCATGTGCAAAATGTGAC 1222
Db 901 AAGACATGTTGTGACCCCATGCTGTTAAACACAAAGACTTGGCCCATGTGCAAAATGTGAC 960
Qy 1223 ATACTCAAGCTTTGGGAATTTGAGTGGATGTTGAAGATGGATCAGTGTCTTTACAAGTC 1282
Db 961 ATACTCAAGCTTTGGGAATTTGAGTGGATGTTGAAGATGGATCAGTGTCTTTACAAGTC 1020
Qy 1283 CCTGTATCCAAATGAATATCTAATAGTGCTCTCCCATGAAGAGATAATCGCAGCGAG 1342
Db 1021 CCTGTATCCAAATGAATATCTAATAGTGCTCTCCCATGAAGAGATAATCGCAGCGAG 1080
Qy 1343 ACCGATCATCTGGATATGCTTCACTAGTACAGGAAACAGATGAACCGCTCTCGAGGAAAC 1402
Db 1081 ACCGATCATCTGGATATGCTTCACTAGTACAGGAAACAGATGAACCGCTCTCGAGGAAAC 1140
Qy 1403 GTGCAGTCAACAAATGAAGTCTACAGCTGTTAAACCATGAGCAAAATCTGTGCGAGTG 1462
Db 1141 GTGCAGTCAACAAATGAAGTCTACAGCTGTTAAACCATGAGCAAAATCTGTGCGAGTG 1200
Qy 1463 GATGTTATTCCTCATGTTGACACCCCAACCTTTGAAGAGACGAACTCCTTAATCAAGAG 1522
Db 1201 GATGTTATTCCTCATGTTGACACCCCAACCTTTGAAGAGACGAACTCCTTAATCAAGAG 1260
Qy 1523 ACTGCTGTTGAGAAATTAATCT 1546
Db 1261 ACTGCTGTTGAGAAATTAATCT 1284

RESULT 6
AAS25884
ID AAS25884 standard; cDNA; 2306 BP.
AC AAS25884;
XX
XX
DT 07-NOV-2001 (first entry)
XX
DE Human cDNA encoding a novel secreted protein, Seq ID 63.
XX
XX Human; immunosuppressive; antiarthritic; ss; antirheumatic;
KW cytostatic; cardiant; vasotropic; cerebroprotective; nootropic;
KW neuroprotective; antibacterial; virucide; fungicide; ophthalmological;
KW vulnary; secreted protein; rheumatoid arthritis;
KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;
KW cerebrovascular disorder; cerebral ischaemia; angiogenesis;
KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;
KW corneal infection; wound healing; epithelial cell proliferation;
KW skin ageing; food additive; preservative; antiproliferative.
XX
OS Homo sapiens.
XX
XX WO20015322-A2.
PN
XX
XX
PD 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01341.
PF
XX 31-JAN-2000; 2000US-0179065.
PR

PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUL-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0232968.
PR 14-SEP-2000; 2000US-0232997.
PR 14-SEP-2000; 2000US-0232998.
PR 14-SEP-2000; 2000US-0232999.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.

PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239335.
PR 13-OCT-2000; 2000US-0239337.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 17-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249219.
PR 17-NOV-2000; 2000US-0249220.
PR 17-NOV-2000; 2000US-0249221.
PR 17-NOV-2000; 2000US-0249224.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
PR XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-488783/53.

XX P-PSDB; AAU15897.

XX New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives -

XX

FS Claim 1; SEQ ID No 63; 980pp; English.

XX

CC The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They
CC are also used in diagnosing a pathological condition or susceptibility
CC to a pathological condition. Antibodies to the proteins can also
CC be used in alleviating symptoms associated with the disorders and in
CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
CC immunosorbant assays (ELISA). Disorders which are diagnosed or treated
CC include autoimmune diseases e.g. rheumatoid arthritis,
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi
CC and ocular disorders e.g. corneal infection, and many other
CC disorders listed in the specification. The polypeptides can also
CC be used to aid wound healing and epithelial cell proliferation, to
CC prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence encodes a novel secreted protein of the invention.

Query Match 66.4%; Score 1178.8; DB 22; Length 2306;

Best Local Similarity 99.0%; Pred. No. 0;

Matches 1217; Conservative 0; Mismatches 8; Indels 4; Gaps 3;

QY 547 CTGTAAACCGGCACAGCAATTTTCAGGTGCCCGTGGGGGAGACCGTGCAGTCTC 606

DB 12 CTGTAAACCGGCACAGCAATTTTCAGGTGCCCGTGGGGGAGACCGTGCAGTCTC 71

QY 607 TTGTTGGCCCTCATCCACCGCGGGGGTGCACCTTCCGACACAGATCCATCTCGC 666

DB 72 TTGTTGGCCCTCATCCACCGCGGGGGTGCACCTTCCGACACAGATCCATCTCGC 131

QY 667 TTATGAGAGATGGCGCTCGGAGCGCTCATCTTTAACTTCCCGGAGCCCAATGAGGT 726

DB 132 TTATGAGAGAGGGGGCTCTGGAGCGCTCATCTTTAACTTCCCGGAGCCCAATGAGGT 191

QY 727 CATCCCATGTCTACCCGGGTGCAGTAGACATTTGTCGAATCATGATCGCAATCTGAA 786

DB 192 CATCCCATGTCTACCCGGGTGCAGTAGACATTTGTCGAATCATGATCGCAATCTGAA 251

QY 787 AGGCACAAAAATTTCTGCAATCTATTCAAAGAGGCATACAAGTGAATGGTCATAGAAGT 846

DB 252 AGGCACAAAAATTTCTGCAATCTATTCAAAGAGGCATACAAGTGAATGGTCATAGAAGT 311

QY 847 AGGGAAAAACAATGGCCCTTGGGTGAATCACTATTCAA - TTTTTCGTTTCTGTGCTCT 905

DB 312 AGGGAAAAACAATGGCCCTTGGGTGAATCACTATTCAAATTTTTCGTTTCTGTGCTCT 371

QY 906 TTTTTCGTTTCTGTGCTCTGTTGCTATTTTATCTTTTATCTCTCGAGGCTTAC 965

DB 372 TTTTTCGTTTCTGTGCTCTGTTGCTATTTTATCTTTTATCTCTCGAGGCTTAC 431

QY 966 GGAATGCAAGAGCTCAAAGCAGGAGCAGAGCAATTAAGCAGATGCTTAAAAAGCTA 1025

DB 432 GGAATGCAAGAGCTCAAAGCAGGAGCAGAGCAATTAAGCAGATGCTTAAAAAGCTA 491

QY 1026 TTGGAAGGCTTCAACTACGCACTGAAACAAGGAGACAAGAAATTCGGCTCTGATGAG 1085

DB 492 TTGGAAGGCTTCAACTACGCACTGAAACAAGGAGACAAGAAATTCGGCTCTGATGAG 551

QY 1086 ATAGTTGCTGTGTGCTGCAATTAATTTGATTAACCAATTAATTTGATCGCATCTTAAGT 1145

DB 552 ATAGTTGCTGTGTGCTGCAATTAATTTGATTAACCAATTAATTTGATCGCATCTTAAGT 611

QY 1146 GCAACCATATTTTCCATAAGACATGTGTTGACCCCATGCTGTTAAACACCAAGACTTGCC 1205

Db 612 GCAACCATATTTTCATAGACATGTGTGACCCATGGCTGTAGAACACAGACTTGC 671
 Qy 1206 CCATGTGCAATGTGACATCTCAAGCTTTGGGAAATGAGGTGGATGTTGAAGATGGAT 1265
 Db 672 CCATGTGCAATGTGACATCTCAAGCTTTGGGAAATGAGGTGGATGTTGAAGATGGAT 731
 Qy 1266 CAGTGTCTTTACAGTCCCTGTATCCCAATGAATATCTATAGTGCCTCCCTCCATGAAG 1325
 Db 732 CAGTGTCTTTACAGTCCCTGTATCCCAATGAATATCTATAGTGCCTCCCTCCATGAAG 791
 Qy 1326 AGGATAATCGCAGGAGACCCATCATCTGGATATGCTTCAGTACAGGGAACAGATGAAC 1385
 Db 792 AGGATAATCGCAGGAGACCCATCATCTGGATATGCTTCAGTACAGGGAACAGATGAAC 851
 Qy 1386 CGCCTCTGGAGGACACGTGAGTCAACAAATGAAGTCTACAGCTGGTAAACCATGAAG 1445
 Db 852 CGCCTCTGGAGGACACGTGAGTCAACAAATGAAGTCTACAGCTGGTAAACCATGAAG 911
 Qy 1446 CAAATTTCTGTGCGAGTGTATTTCTCATGTNGACAACCCCAACCTTTGAAGAAGACG 1505
 Db 912 CAAATTTCTGTGCGAGTGTATTTCTCATGTNGACAACCCCAACCTTTGAAGAAGACG 971
 Qy 1506 AAATCTCTTAATCAAGAGACTGTGTTCGAGAAATTAATCTTAAATCTGTGTAATAGA 1565
 Db 972 AAATCTCTTAATCAAGAGACTGTGTTCGAGAAATTAATCTTAAATCTGTGTAATAGA 1031
 Qy 1566 AAATCTCTTAATCAAGAGACTGTGTTCGAGAAATTAATCTTAAATCTGTGTAATAGA 1625
 Db 1032 AAATCTCTTAATCAAGAGACTGTGTTCGAGAAATTAATCTTAAATCTGTGTAATAGA 1091
 Qy 1626 TGGATAAATTTAATAAATAAGAGTATCTGAAAGTGTCTCAGATGACTAATATTATGCT 1685
 Db 1092 TGGATAAATTTAATAAATAAGAGTATCTGAAAGTGTCTCAGATGACTAATATTATGCT 1151
 Qy 1686 ATAGTAAATGGCTTAAATAATATTAACTGTAACTTTTTCACCAAACTCATTATA 1745
 Db 1152 ATAGTAAAT--TGGCTTAAATAATATTAACTGTAACTTTTTCACCAAACTCATTATA 1208
 Qy 1746 ATATTTTTCATAGGCAAGTTCCTCTCAG 1774
 Db 1209 ATATTTTTCATAGGCAAGTTCCTCTCAG 1237

RESULT 7

ABL90796

ID ABL90796 standard; cDNA; 1249 BP.

XX AC ABL90796;

XX XX

XX DT 24-MAY-2002 (first entry)

XX DE Human polynucleotide SEQ ID NO 1358.

XX KW Cytostatic; immunosuppressive; nontoxic; neuroprotective; antiviral;

XX KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;

XX KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;

XX KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;

XX KW neurological disease; infection; human; secreted protein; gene; ss.

XX OS Homo sapiens.

XX PN WO200190304-A2.

XX XX 29-NOV-2001.

XX XX 18-MAY-2001; 2001WO-US16450.

XX XX 19-MAY-2000; 2000US-205515P.

XX XX (HUMA-) HUMAN GENOME SCI INC.

XX PI Birse CE, Rosen CA;

XX

DR

XX

PT

PT

PT

PT

XX

PS

XX

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

WPI; 2002-122018/16.

P-P5DB; ABB90387.

Novel 1405 isolated polypeptides, useful for diagnosis, treatment and prevention of neural, immune system, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal and proliferative disorders -

Claim 4; SEQ ID NO 1358; 2081pp + Sequence Listing; English.

The invention relates to novel genes (ABL89449-ABL90853) and proteins (ABB9040-ABB90444) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pat_sequences.

Sequence 1249 BP; 379 A; 258 C; 273 G; 335 T; 4 other;

Query Match

Best Local Similarity 65.8%; Score 1167.2; DB 24; Length 1249;

Matches 1206; Conservative 1; Mismatches 11; Indels 4; Gaps 3

QY 547 CTGTAAACCCGCAACAGATTTTCCAGTGCCTTGGGGAAGACCGTCAAGTCTC 606

Db 18 CTGTAAACCCGCAACAGATTTTCCAGTGCCTTGGGGAAGACCGTCAAGTCTC 77

QY 607 TTGGTTGGCCCTCATCCACGGCGGGCTGACCTTCGCAGACAAGATCCATCTGC 666

Db 78 TTGGTTGGCCCTCATCCACGGCGGGCTGACCTTCGCAGACAAGATCCATCTGC 137

QY 667 TTATGAGAGATGGGCTCTGGAGCCCTCATCTTTAACTTCCCGGACCCGCAATGAGT 726

Db 138 TTATGAGAGATGGGCTCTGGAGCCCTCATCTTTAACTTCCCGGACCCGCAATGAGT 197

QY 727 CATCCCATGTCTCACCCGGGTGCAGTAGACATTTTGCATCATGCGCAATCTGAA 786

Db 198 CATCCCATGTCTCACCCGGGTGCAGTAGACATTTTGCATCATGCGCAATCTGAA 257

QY 787 AGGCACAAAAATTTCTGCAATCTATTCAAGAGGCATACAAGTGACAATGGTCAAGAAGT 846

Db 258 AGGCACAAAAATTTCTGCAATCTATTCAAGAGGCATACAAGTGACAATGGTCAAGAAGT 317

QY 847 AGGAAAAAACAATGGCCCTTGGGTGAACTCACTATTCAATTTTTTCTGTGTCTT 906

Db 318 AGGAAAAAACAATGGCCCTTGGGTGAACTCACTATTCAATTTTTTCTGTGTCTT 377

QY 907 TTTTATTATTACGGGGCAACTGTGGCTATTATTATCTTTTATCTGTCTGCAAGGCTACG 966

Db 378 TTTTATTATTACGGGGCAACTGTGGCTATTATTATCTTTTATCTGTCTGCAAGGCTACG 437

QY 967 GAATGCAAGAGCTCAAGAGGAGAGAGGCAATTTAAAGGAGAGATGCTAAAAAGCTAT 1026

Db 438 GAATGCAAGAGCTCAAGAGGAGAGAGGCAATTTAAAGGAGAGATGCTAAAAAGCTAT 497

QY 1027 TGGAGGCTTCAACTAGGCACACTGAAACAAGAGAGAGAGG- AATTTGGCCCTGATGGAG 1085

Db 498 TGGAGGCTTCAACTAGGCACACTGAAACAAGAGAGAGAGG- AATTTGGCCCTGATGGAG 557

QY 1086 ATAGTTGTGTGTGCTGCAATTGAATTGTATAAAACCAAAATGATTTGGTACGCACTTAAACGT 1145

Db 584 TGAAGCCGCTCTCCGGGCTCTGGTACCGCCGACGGGCCCGGGCGCTCAACGCCCTGTA 643
Qy 552 ACCCGCACACGAATTTACGGTCCCAAGGTTTGGGGAAGCAGCGTCAAGTCTCTTGGT 611
Db 644 ACCCGCACACGAATTTACGGTCCCAAGGTTTGGGGAAGCAGCGTCAAGTCTCTTGGT 703
Qy 612 TGGCCCTCATCAACGGGGGGGGCTGCACCTTTGCGCAGACAAGATCCATCTGCTTATG 671
Db 704 TGGCCCTCATCAACGGGGGGGGCTGCACCTTTGCGCAGACAAGATCCATCTGCTTATG 763
Qy 672 AGAGATGGGCGCTCTGGAGCGGCTCATCTTTAACTTTCCCGGGGACCCGCAATGAGGTATCC 731
Db 764 AGAGAGGGGCTCTGGAGCGGCTCATCTTTAACTTTCCCGGGGACCCGCAATGAGGTATCC 823
Qy 732 CCATGCTCACCCGGGTCAGTACACATTTGTCAAATCATGATCGGCAATCTGAAAGGCA 791
Db 824 CCATGCTCACCCGGGTCAGTACACATTTGTCAAATCATGATTTGGCAATCTGAAAGGAA 883
Qy 792 CAAAAATTCGAATCTATTCAAGAGGACATCAAGTGACAAATGGTTCATAGAGTAGGGA 851
Db 884 CAAAAATTCGAGTCTATTCAAGAGGACATCCCAAGTCAAAATGGTTCATCGAAGTAGGGA 943
Qy 852 AAAACATGGCCCTTGGGTGAATCACTATTCAATTTTTCGTTCTGTCCTTTTCA 911
Db 944 AAAACATGGCCCTTGGGTGAATCACTATTCAATTTTTCGTTCTGTCCTTTTCA 1003
Qy 912 TTATTACGGCGCAACTGTGGCTATTATTATCTTTTATCTGCTCGAAGGCTACGGAATG 971
Db 1004 TAAATTACGGGCAACCGTGGCTATTTCATCTTTTATCTGCTCGAAGATTACGAATG 1063
Qy 972 CAAGAGCTCAAGCAGGAGCAGAGGCAATTTAAAGGCGAGATGCTTAAAGAGCTATTGGAA 1031
Db 1064 CAAGAGCTCAAGCAGGAGCAGAGGCAATTTAAAGGCGAGATGCTTAAAGAGCTATTGGAA 1123
Qy 1032 GGCCTTCAACTAGCACA CTGAAACAAGGAGACAAGGAATTTGGCCCTGATGGAGATAGTT 1091
Db 1124 AGCTTCAGTGGCCACTTTGAAACAGGAGACAAGGAATTTGGCCCTGATGGAGATAGTT 1183
Qy 1092 GTGCTGTGTGCAATTTGATATAACCAAAATGATTTGGTACGATCTTAAACGTGCAACC 1151
Db 1184 GTGCTGTGTGCAATTTGATATAACCAAAATGATTTGGTACGATCTTAAACGTGCAACC 1243
Qy 1152 ATATTTTCATAGACATGCTGTGACCATGGCTGTTTAAACACACAGACTTGGCCCATGT 1211
Db 1244 ATATTTTCATAGACATGCTGTGACCATGGCTGTTTAAACACACAGACTTGGCCCATGT 1303
Qy 1212 GCAATGTGACATCTCAAGCTTTTGGGAATTTGAGTGGATGTTTGAAGATCGATCAGTGT 1271
Db 1304 GCAAGTGTGACATCTCAAGCTTTTGGGAATTTGAGTGGATGTTTGAAGATCGATCAGTGT 1363
Qy 1272 CTTTCAAGTCCCTGTATCCAAATGAATATCTAAATGATGCTCTCCCTCCCATGAAGAGATA 1331
Db 1364 CTTTCAAGTCCCTGTATCCAAATGAATATCTAAATGATGCTCTCCCTCCCATGAAGAGATA 1423
Qy 1332 ATCGCAGCAGACCGCATCATCTGATATGCTTTCAGTACAGGGAACAGATGAACCGCTC 1391
Db 1424 GTCCGAGTGAAGTGCATCATCTGATATGCTTTCAGTACAGGGAACAGATGAACCGCTC 1483
Qy 1392 TGGAGAAACCGTGACATCAACAAATGAAGTCTACAGTGGTAAACCATGAAGCAAT 1451
Db 1484 TGGAGAAACCGTGACATCAACAAATGAAGTCTACAGTGGTAAACCATGAAGCAAT 1543
Qy 1452 CTGTGCACTGGATGTTATCTCTCATGTTGACAAACCCACCTTTGAAAGAGACGAACTC 1511
Db 1544 CTGTGCACTGGATGTTATCTCTCATGTTGACAAACCCACCTTTGAAAGAGATGAACCTC 1603
Qy 1512 CTAATCAAGAGACTGCTGTTTCGAGAAATTAATCTTT - AAAATCTGTGTAATAGAAAAT 1570
Db 1604 CTGATCAAGAGGAGCTGTTTCGGGAGATTAATCTTTAAATCTGTGCAATAGAAAAT 1663
Qy 1571 TGAACCAATTAG - TAATAACAGAACTGCCAATCAGGGCCTAGTTT - CTATTATAAATTTGG 1628
Db 1664 TGAACCGTTAGTTTAAACAGGAGCTGCCAATCAGGGCCTAGTTTACTATGATGAATGAACTGG 1723

Qy 1629 ATAAATTTAATAAATAAGAGTGATGACTGAAGTGCTCAGATGACTAATATTATGCTATA 1688
Db 1724 GTAAA ---CGTAAACAAGAATGATGACTGAAGTGCTGAGTAACTTATATTACTATA 1780
Qy 1689 GTTAAATGGCTTAAAAATATTTAACTGTTTAACTTTTCCCAAACTCAATTATATA 1748
Db 1781 GTTAAA ---TGGCTTAAACATATTTACCCAGTACCGTTTTT---CCACAAACTCACCATAACG 1836
Qy 1749 TTTTTCATAGGCAAGTTTCCCTCT 1771
Db 1837 TTTTTCATAGGCAAGTTTCCCTCT 1859
RESULT 9
AAS26340
ID AAS26340 standard; cDNA; 1250 BP.
XX
AC AAS26340;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human cDNA encoding a novel secreted protein, Seq ID 519.
XX
KW Human; immunosuppressive; antiarthritic; ss; antirheumatic;
cytostatic; cardiant; vasotropic; cerebroprotective; nootropic;
neuroprotective; antibacterial; virucide; fungicide; ophthalmological;
vulnerable; secreted protein; rheumatoid arthritis;
hyperproliferative disorder; cardiovascular disorder; cardiac arrest;
cerebrovascular disorder; cerebral ischaemia; angiogenesis;
nervous system disorder; Alzheimer's disease; infection; ocular disorder;
corneal infection; wound healing; epithelial cell proliferation;
skin ageing; food additive; preservative; antiproliferative.
XX
OS Homo sapiens.
XX
PN WO200155322-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01341.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-019874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.

PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226686.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0228287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0234999.
 PR 27-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 29-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246533.
 PR 08-NOV-2000; 2000US-0246534.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.

PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 08-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX XX
 (HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-488783/53.
 P-PSDB; AAU16353.

New nucleic acid molecules encoding 461 human secreted proteins for diagnosing, preventing, treating or ameliorating medical conditions and used as food additives or preservatives -

Claim 1; SEQ ID No 519; 980pp; English.

The invention relates to isolated nucleic acid molecules and their encoded secreted proteins. The nucleic acids and proteins are used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used in diagnosing a pathological condition or susceptibility to a pathological condition. Antibodies to the proteins can also be used in alleviating symptoms associated with the disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme linked immunosorbant assays (ELISA). Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection, and many other disorders listed in the specification. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors and other nutritional components. The present sequence encodes a novel secreted protein of the invention.

Query Match 65.3%; Score 1157.6; DB 22; Length 1250;
 Best Local Similarity 98.8%; Pred. No. 8.9e-312;

Matches 1208; Conservative 0; Mismatches 10; Indels 5; Gaps 4;

Qy	547	CTGTAAACCGGCACAGAAATTTTACGGTGGCCACCGTTTGGGGAGCAACCGTGCAGTCTC	606
Db	18	CTGTAAACCGGCACAGAAATTTTACGGTGGCCACCGTTTGGGGAGCAACCGTGCAGTCTC	77
Qy	607	TTGGTTGGCCCTCATCCAAACGCGCGGGGGCTGCACCTTTCGCAGACAAGATCCATCTGGC	666
Db	78	TTGGTTGGCCCTCATCCAAACGCGCGGGGGCTGCACCTTTCGCAGACAAGATCCATCTGGC	137
Qy	667	TTATGAGAGATGGGCGTCTGGAGCGCTCATCTTTAACTTCCCGGGACCCGCAATGAGGT	726
Db	138	TTATGAGAGAGGGGGCTGGAGCGCTCATCTTTAACTTCCCGGGACCCGCAATGAGGT	197
Qy	727	CATCCCCCATGTCTCACCCGGGTGCAGTAGACATTTGTGCAATCATGATCGCAATCTGAA	786
Db	198	CATCCCCCATGTCTCACCCGGGTGCAGTAGACATTTGTGCAATCATGATCGCAATCTGAA	257
Qy	787	AGGCACAAAAATCTGCAATCTATTCAAAGAGGCATACAAGTGAACAATGGTCATAGAAGT	846
Db	258	AGGCACAAAAATCTGCAATCTATTCAAAGAGGCATACAAGTGAACAATGGTCATAGAAGT	317
Qy	847	AGGAAAAAACAATGGGCCCTTGGGTGAATCACTATTTCAA- TTTTTCCTTCTGTGTCTCT	905
Db	318	AGGAAAAAACAATGGGCCCTTGGGTGAATCACTATTTCAA- TTTTTCCTTCTGTGTCTCT	377
Qy	906	TTTTTATTTATACGGCGCAACTGTGGGCTATTTTATCTTTTATTTCTGCTCGAAGGCTAC	965
Db	378	TTTTTATTTATACGGCGCAACTGTGGGCTATTTTATCTTTTATTTCTGCTCGAAGGCTAC	437
Qy	966	GGATGCAGAGCTCAAGACGAGGACGAGGCAATTAAGCGAGATGCTTAAAAAGCTA	1025
Db	438	GGATGCAGAGCTCAAGACGAGGACGAGGCAATTAAGCGAGATGCTTAAAAAGCTA	497
Qy	1026	TTGGAAGGCTTCAACTACGCACACTGAAACAAAGGAGACAAG- AAATTGGCCCTGATGGA	1084
Db	498	TTGGAAGGCTTCAACTACGCACACTGAAACAAAGGAGACAAG- AAATTGGCCCTGATGGA	557
Qy	1085	GATAGTGTGCTGTGTGCATTCGAATGTATTAACCAAAATGATTTGGTACGCATCTTAACG	1144
Db	558	GATAGTGTGCTGTGTGCATTCGAATGTATTAACCAAAATGATTTGGTACGCATCTTAACG	617
Qy	1145	TGCAACCATATTTTCCATAAGACATGTGTGACCCATGGCTGTTTAAACACAAAGCTTGC	1204
Db	618	TGCAACCATATTTTCCATAAGACATGTGTGACCCATGGCTGTTTAAACACAAAGCTTGC	677
Qy	1205	CCCATGTGCAAAATGTACACATCTCAAGCTTTGGGAATTGAGGTGGATGTTGAAGATGGA	1264
Db	678	CCCATGTGCAAAATGTACACATCTCAAGCTTTGGGAATTGAGGTGGATGTTGAAGATGGA	737
Qy	1265	TCAGTGTCTTTTACAAGTCCCTGTATCCAATGAAATATCTAATAGTGCCTCTCCCATGAA	1324
Db	738	TCAGTGTCTTTTACAAGTCCCTGTATCCAATGAAATATCTAATAGTGCCTCTCCCATGAA	797
Qy	1325	GAGATTAATCGACGAGACCGCATCATCTGATATGCTTCAGTACAGGGAACAGATGAA	1384
Db	798	GAGATTAATCGACGAGACCGCATCATCTGATATGCTTCAGTACAGGGAACAGATGAA	857
Qy	1385	CCGCTCTCGAGGAAACACGTGCAGTCAAACAAATGAAAGTCTACAGCTGGTAAACATGAA	1444
Db	858	CCGCTCTCGAGGAAACACGTGCAGTCAAACAAATGAAAGTCTACAGCTGGTAAACATGAA	917
Qy	1445	GCAAAATCTGTGGAGTGGATGTTATTTCTCATGTNGACAACCCAACTTTTGAAGAGAC	1504
Db	918	GCAAAATCTGTGGAGTGGATGTTATTTCTCATGTGTGAACCCAACTTTTGAAGAGAC	977
Qy	1505	GAAACTCCTAATCAAGAGACTGCTGTTCAGAGAAATTAATCTTAAATCTGTGTAATATAG	1564
Db	978	GAAACTCCTAATCAAGAGACTGCTGTTCAGAGAAATTAATCTTAAATCTGTGTAATATAG	1037
Qy	1565	AAAACTTGAAACCAATTAGTAATAACAGAACTGCCAAATCAGGGCTAGTTTCTATTATAAA	1624
Db	1038	AAAACTTGAAACCAATTAGTAATAACAGAACTGCCAAATCAGGGCTAGTTTCTATTATAAA	1097

Qy	1625	TTCGGATAAAATTTAAATAAAATAGAGTGATAC	CTGAAAGTGCTCAGATGACCTAAATATTATGC	1684
Db	1098	TTCGGATAAAATTTAAATAAAATAGAGTGATAC	CTGAAAGTGCTCAGATGACCTAAATATTATGC	1157
Qy	1685	TATAGTTAAATGCGCTTAAATAATATTTAACTG	TTTTTTCACCAAACTCATTAT	1744
Db	1158	TATAGTAAAA--TGGCTTAAATAATTTAACTGT	TAACTTTTTTTCCTCAAACTCATTAT	1214
Qy	1745	AAATATTTTTCATAGGCAAGTTTC	1767	
Db	1215	AAATATTTTTCATAGGCAATTTCC	1237	

RESULT	10
ABK12983	
ID	ABK12983 standard; cDNA; 1287 bp.
XX	
XX	ABK12983;
XX	.
DT	23-APR-2002 (first entry)
XX	
DE	Mouse goliath cDNA sequence.
XX	
KW	Mouse; goliath protein; antiangiogenic; vasotropi
KW	doseage form; angiogenesis; neurogenesis; tumour;
KW	cancer; ischaemia; neuroblastoma; neurodegenerati
KW	unrecovered nerve trauma; gene; ss.
XX	
OS	Mus sp.
XX	
XX	
FH	Key Location/Qualifiers
CDS	1..1287
FT	/tag= a
FT	/product= "Mouse.goliath.protein"
FT	

Sequence 1287 BP; 308 A; 320 C; 359 G; 300 T; 0 other;

Query Match 62.4%; Score 1106.8; DB 24; Length 1287;
 Best Local Similarity 91.2%; Pred. No. 1.3e-297;
 Matches 1174; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

QY 263 ATGGGGCCGGCTGGGGCGGGGCTCTCTGCGCGGGTGGCTGGCGCTTTTCCAGATGG 322
 DB 1 ATGGGGCCGGCGGGGATCGGGGTCTACTGCGCGGGCGGTGCGGAGCTGCCCGGTA 60

QY 323 CTGGCATGGTGTCTCTGCTGGCCCTGAGTCCGACAGCCACCGGTTCCGCGGGGGCTGAA 382
 DB 61 CTGGCTGGTGTCTCTCTGCTGGCTGAGTCCGACAGCCGCGGTTCCCGGGAGCCGAA 120

QY 383 GCAGTGTGGACCGGTACCTCAACGTGTCTGCGGGGTTCCGCAACAGGGAGTGAACCGT 442
 DB 121 GCGTGTGGACCGGTACCTCAACGTGTCTGCGGGGTTCCGCAACAGGGAGTGAACCGC 180

QY 443 ACGTGTGGAGCTGAGCGAGGAGGGCGTGTACGGCCAGCACTCGCGCGTGGAGCGTGTG 502
 DB 181 ACGTGTGGAGCTGAGCGAGGAGGGCGTGTACGGCCAGCACTCGCGCGTGGAGCGCGT 240

QY 503 CTGGGGTCTGTGATCGCGCCGACGGGCGCGGGCGCTTAAACGCTGTAAACCGCACAG 562
 DB 241 TCGGGGTCTGTGATCGCGCCGACGGGCGCGGGCGCTCAACGCTGTAAACCGCACAC 300

QY 563 AATTTCACGGTGCACGGTTTGGGAGGACACCGTGCAGTCTCTTGGTTGGCCCTCATC 622
 DB 301 AATTTCACGGTGCACGGTTTGGGAGGACACCGTGCAGTCTCTTGGTTGGCCCTCATC 360

QY 623 CAACGGCGCGGGCTGCACCTTCGACAGCAAGATCCATCTGGCTTATGAGAGATGGGG 682
 DB 361 CAACGGCGTGGAGCTGCACCTTCGCGGACAGATCCATCTGGCTTACAGAGAGGGCT 420

QY 683 TCTGGAGCGTGTATCTTTAACTTCCCGGAGACCGCAATGAGTGTATCCCATGTCTAC 742
 DB 421 TCTGGAGCGTGTATCTTTAACTTCCCGGAGACCGCAATGAGTGTATCCCATGTCTAC 480

QY 743 CCGGGTGCAGTACACATCTTGTCAATCATGATCGGCAATCTGAAAGGCACAAATTCG 802
 DB 481 CCGGGTGTGGAGCATTTGTGCAATCATGATGGCAATCTGAAAGGCACAAATTCG 540

QY 803 CAATCTATTCAAGAGGCATACAAAGTGAATGGTGCATAGAGTAGGGGAAAAAATGCG 862
 DB 541 CAGTCTATTCAAGAGGCATCAAGTCAATGGTGCATAGAGTAGGGGAAAAAATGCG 600

QY 863 CTTGGGTGAATCACTATCAATTTTTCGTTTCTGTGCTCTTTTATTAATACGGCG 922
 DB 601 CTTGGGTGAATCACTATCAATTTTTCGTTTCTGTGCTCTTTTATTAATACGGCA 660

QY 923 GCAACTGTGGCTATTTATCTTTTATCTGCTCGAAGGCTACGGAATCAAGAGCTCAA 982
 DB 661 GCAACGTGGCTATTTATCTTTTATCTGCTCGAAGATTAAGAAATGCAAGAGCTCAA 720

QY 983 AGCAGGAAGCAGAGGCAATTAAGGCGAGATGCTTAAAAAGCTTATTCGAAAGGCTTCACTA 1042
 DB 721 AGCAGGAAGCAGAGGCAATTAAGGCGAGATGCTTAAAAAGCTTATTCGAAAGCTTCACTG 780

QY 1043 CGCACACTGAAACAAGGAGCAAGGAATTTGGCCCTGATGGAGATAGTGTGCTGTGTC 1102
 DB 781 CGCACCTTGAACAAGGAGCAAGGAATTTGGCCCTGATGGAGATAGTGTGCTGTGTC 840

QY 1103 ATTGAATTTGATAAACCAGATTTTGGTACGATCTTAACTGCGCAACCATATTTTCCAT 1162
 DB 841 ATTGAGCTCTAAGCCAAATGATTTGGTGGCATCTTAACCTGGCATCATATTTTCCAT 900

QY 1163 AAGACATGTGTGACCCCATGGCTTTAAACAACAAGCTTTGCCCATGTGCAAAATGTGAC 1222
 DB 901 AAGACATGTGTGACCCCATGGCTTTAAACAACAAGCTTTGCCCATGTGCAAAATGTGAC 960

QY 1223 ATACTCAAGCTTTGGGAATTTGAGTGGATGTTGAAGATGGATCAGTGTCTTTACAGTC 1282
 DB 961 ATCTCAAGCTCTGGGAATTTGAGTGGATGTTGAAGATGGATCAGTGTCTTTACAGTT 1020

QY 1283 CCTGTATCCAAATGAATATCTAATAGTGCCTCTCCCATGAAGAGGATATCGCAGCGAG 1342
 DB 1021 CCTGTATCCAAATGAAGCATCTAATAGTGCCTCTCCCATGAAGAGGACAGTCGCAGTGAG 1080

QY 1343 ACCGCATCATCTGATATGCTTCACTACAGGAAACAGATGAACCGCCTCTGGAGGAACAC 1402
 DB 1081 ACTGCATCATCTGATATGCTTCACTACAGGAGGAGATGAGCCACCTCTGGAGGAACAT 1140

QY 1403 GTGCAGTCAACAAATGAAAGTCTACAGCTGTGTAAACCATGAAGCAAAATCTGTGGCAGTG 1462
 DB 1141 GCGCAGTCAGCAAAATGAAATCTACAGCTGTGTAAACCATGAAGCAAAATCTGTGGCAGTG 1200

QY 1463 GATGTATCTCTCATGTGACCAACCAACCTTTGAAGAACAGCAAACTCTTAATCAAGAG 1522
 DB 1201 GATGTGTGTCCTCATGTGTGAACCAACCAACCTTTGAAGAACAGCAAACTCTTAATCAAGAG 1260

QY 1523 ACTGCTGTTCGAGAAATTAATCTTAA 1549
 DB 1261 GCAGCTGTTCGGGAGATTAATCTTAA 1287

RESULT 11

ABK12992
 ID ABK12992 standard; DNA; 1287 BP.
 XX
 AC ABK12992;
 XX
 DT 23-APR-2002 (first entry)
 XX
 DE Synthetic goliath DNA sequence #6.
 XX
 KW Goliath protein; antiangiogenic; vasotropic; gene therapy;
 KW dosage form; angiogenesis; neurogenesis; tumour; vascularisation;
 KW cancer; ischaemia; neuroblastoma; neurodegenerative disease;
 KW unrecovered nerve trauma; ds.
 XX
 OS Synthetic.
 XX
 PN WO200193681-A1.
 XX
 PD 13-DEC-2001.
 XX
 PP 01-JUN-2001; 2001WO-US18000.
 XX
 PR 02-JUN-2000; 2000US-0586398.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Harland R, Baker JC;
 XX
 DR WPI; 2002-147637/19.
 XX
 PT New compositions comprising goliath proteins, useful for modulating
 PT angiogenesis or neurogenesis in mammals e.g. for preventing or
 PT treating undesirable vascularisation of a tumour, ischaemia or
 PT neurodegenerative disease
 XX
 FS Disclosure; Page 42; 45pp; English.
 XX
 CC The present invention relates to a new pharmaceutical composition that
 CC comprises a goliath polypeptide in dosage form. The goliath polypeptide
 CC has a sequence identity of at least 75% to the protein sequences
 CC (AAU74918-AAU74921) fully defined in the specification. The composition
 CC is useful for modulating angiogenesis or neurogenesis in mammals,
 CC particularly in humans or mice. Specifically, the composition is useful
 CC for the prophylactic and/or therapeutic treatment of excess angiogenesis
 CC e.g. undesirable vascularisation of a tumour or insufficient angiogenesis
 CC e.g. ischaemia, excess neurogenesis e.g. neuroblastoma or insufficient
 CC neurogenesis e.g. neurodegenerative disease or unrecovered nerve trauma.
 CC The present nucleic acid sequence represents synthetic DNA sequence #6
 CC that is one of several artificial goliath nucleic acids (ABK12987-
 CC ABK12992) used in the invention for modulating angiogenesis or
 CC neurogenesis.

```
XX SQ Sequence 1287 BP; 312 A; 321 C; 357 G; 297 T; 0 other;
Query Match 59.3%; Score 1052.4; DB 24; Length 1287;
Best Local Similarity 88.6%; Pred. No. 1.8e-282;
Matches 1140; Conservative 0; Mismatches 147; Indels 0; Gaps 0;
QY 263 ATGGGGCGCGCGCTGGGGCGGGGTCTCTCGCGCGGGTGGCTGGCGCTTTTCCAGATTG 322
DB 1 ATGGGGCGCGCGCGGGATCGGGTCTACTGCGCGGGGGCTGCGAGCTGCCGACTA 60
QY 323 CTGGCATGTGTCTTCTGCTGGCCCTGAGTCCGGAGGCAACCCGGTTCCGGGGGGCTGAA 382
DB 61 CTGGCTTGGTGTCTTCTTCTGGCTCTGAGTCCCGACCGCGCGGTTCCCGGGGAGCGAA 120
QY 383 GCAGTGTGGACCGGTACTCAAGTGTCTGGCGGGTCCGACACGGGAGTGAACCGT 442
DB 121 GCCGTGTGAACCGGTACTCAAGTGTCTGGCGGGTCCGACACCGGAGTGAACCGC 180
QY 443 ACGGTGTGGAGCTGAGCGAGGCGGTGTACGGCCAGGACTCGCGCTGGAGCCCTGTG 502
DB 181 ACGGTGTGGAGCTGAGCGAGGCGGTGTACGGCCAGGACTCGCGCTGGAGCCCGTC 240
QY 503 GCTGGGTCTGCTGTAACCGCGGACGGGCGCGGGCGCTTAACGGCTGTAAACCGCACAG 562
DB 241 TCCGGGGTCTGTTGGTCCCGCGGACGGGCGCGGAGCGCTCAACGCTGTAAATCCGCAACC 300
QY 563 AATTTCACGGTCCCGACCGTTTGGGGAAGCACCGTCAAGTCTCTTGGTTGGCCCTCATC 622
DB 301 AATTTCACGGTCCCGACCGTTTGGGGAAGCACCGTCAAGTCTCTTGGTTGGCCCTCATC 360
QY 623 CAACCGCGCGGGCTGCACCTTCGCGAGACAAGATCCATCTGGCTTATGAGAGATGGCG 682
DB 361 CATCGCGTGGAGCTGCACCTTCGCGGACAAGATCCAGCTGGCTTCAGAGAGAGAGCT 420
QY 683 TCTGGAGCGCTCATCTTTAACTTCCCGGACCGCGATGAGTGTCACTCCCGCTGTCTCAC 742
DB 421 TCTGGAGCGCTCATCTTTAACTTCCCGGACCGCGATGAGTGTCACTCCCGCTGTCTCAC 480
QY 743 CCGGCTGAGTGTGATCTTTGCAATCATGATCGGCAATCTGAAAGGCACAAAAATCTG 802
DB 481 CCGGCTGAGTGTGATCTTTGCAATCATGATCGGCAATCTGAAAGGCACAAAAATCTG 540
QY 803 CAATCTATTCAAGAGGATCAAGTGTGATGATGATGATGATGATGATGATGATGATGATG 862
DB 541 CAGTCTATTCAAGAGGATCAAGTGTGATGATGATGATGATGATGATGATGATGATGATG 600
QY 863 CTTGGGTGAATCACTATTCAATTTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTG 922
DB 601 CTTGGGTGAATCACTATTCAATTTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTG 660
QY 923 GCAATGTGGGTATTTTATCTTTTATCTGCTGAAGGCTACGGAATGCAAGAGCTCAA 982
DB 661 GCAACCGTGGGTATTTTATCTTTTATCTGCTGAAGGCTACGGAATGCAAGAGCTCAA 720
QY 983 AGCAGGAGGAGGAGGATTAAGGAGGAGGATTAAGGAGGATTAAGGAGGATTAAGGAGG 1042
DB 721 AGTAGAAGGAGGAGGATTAAGGAGGAGGATTAAGGAGGATTAAGGAGGATTAAGGAGG 780
QY 1043 CGCACACTGAAACAAAGGAGACAAAGGAAATGGCCCTGATGAGATAGTGTGCTGTGC 1102
DB 781 CGCACCTTGAACAAAGGAGACAAAGGAAATGGCCCTGATGAGATAGTGTGCTGTGC 840
QY 1103 ATTGAATGTATTAACCAAGATGATTTGATGATGATGATGATGATGATGATGATGATGAT 1162
DB 841 ATTGAGCTATTAAGCAAAATGATTTGATGATGATGATGATGATGATGATGATGATGAT 900
QY 1163 AAGACATGTGTGACCCATGCTGTGTTAAACACAGAGCTGCCCCATGTGCAATGTGAC 1222
DB 901 AAGACATGTGTGACCCATGCTGTGTTAAACACAGAGCTTGGCCCATGTGCAATGTGAC 960
QY 1223 ATACTCAAGCTTTGGGAATTTGAGTGTGATGTTGAAGATGATGATGATGATGATGATGAT 1282
DB 1223 ATACTCAAGCTTTGGGAATTTGAGTGTGATGTTGAAGATGATGATGATGATGATGATGAT
```

RESULT 12

ABK12990

ID ABK12990 standard; DNA; 1287 BP.

XX ABK12990;

XX 23-APR-2002 (first entry)

XX Synthetic goliath DNA sequence #4.

Goliath protein; antiangiogenic; vasotropic; gene therapy;
dosage form; angiogenesis; neurogenesis; tumour; vascularisation;
cancer; ischaemia; neuroblastoma; neurodegenerative disease;
unrecovered nerve trauma; ds.

XX Synthetic.

XX WO200193681-A1.

XX 13-DEC-2001.

XX 01-JUN-2001; 2001WO-US18000.

XX 02-JUN-2000; 2000US-0586398.

XX (REGC) UNIV CALIFORNIA.

XX Harland R, Baker JC;

XX WPI; 2002-147637/19.

New compositions comprising goliath proteins, useful for modulating
angiogenesis or neurogenesis in mammals e.g. for preventing or
treating undesirable vascularisation of a tumour, ischaemia or
neurodegenerative disease

XX Disclosure; Page 41; 45pp; English.

The present invention relates to a new pharmaceutical composition that
comprises a goliath polypeptide in dosage form. The goliath polypeptide
has a sequence identity of at least 75% to the protein sequences
(AAU74918-AAU74921) fully defined in the specification. The composition
is useful for modulating angiogenesis or neurogenesis in mammals,
particularly in humans or mice. Specifically, the composition is useful
for the prophylactic and/or therapeutic treatment of excess angiogenesis
e.g. undesirable vascularisation of a tumour or insufficient angiogenesis
e.g. ischaemia, excess neurogenesis e.g. neuroblastoma or insufficient
neurogenesis e.g. neurodegenerative disease or unrecovered nerve trauma.
The present nucleic acid sequence represents synthetic DNA sequence #4
that is one of several artificial goliath nucleic acids (ABK12997-

CC ABK12992) used in the invention for modulating angiogenesis or
 CC neurogenesis.
 XX
 SQ Sequence 1287 BP; 312 A; 321 C; 356 G; 298 T; 0 other;
 Query Match 58.8%; Score 1042.8; DB 24; Length 1287;
 Best Local Similarity 88.1%; Pred. No. 8.7e-280;
 Matches 1134; Conservative 0; Mismatches 153; Indels 0; Gaps 0;

QY 263 ATGGGGCCGCGCTGGGGCCGGGGTCTCTCGCCGGGTGGTGGCGCTTTTCCAGATTG 322
 Db 1 ATGGGGCCGCGCGGGATAGGGTCTACTGCGCGGTGGCTGCGAGCTGCCGCCCTA 60

QY 323 CTGCGATGGTCTCTGCTGGCCCTGAGTCCGACGACGCCCGGTTCGGGGGGCTGAA 382
 Db 61 CTGCGTGGGTCTGCTTCTGGCTCTGAGTCCACACGCGCCGGTTCGGGTGGAGCCGAA 120

QY 383 GCAGTGTGGACCGGCTACCTCAACGCTGCTCTGGCGGGTTCGGCACACGGGAGTGAACCGT 442
 Db 121 GCGGTGTGACCGGTACTCAACGCTGCTCTGGCGGGTTCGGCAACCGGAGTGAACCGC 180

QY 443 ACGGTGTGGAGCTGAGCGAGGAGGGGTGTTACGGCCAGACTTCGCCGTGGAGCTGTG 502
 Db 181 ACTGTGTGGAGCTGAGCGAGGAGGGGTGTACGGCCAGACTTCGCCGTGGAGCCAGTC 240

QY 503 GCTGGGGTCTGCTACCGCCGACGGGCCCGGGCGCTTAAACGCTGTAAACCGCACAG 562
 Db 241 TCGGGGTCTGCTTCCGCCCGACGGGCCCGGGCGCTTAAACGCTGTAAACCGCACAG 300

QY 563 AATTTCACGGTGGCCCGGTTTGGGGAAGCACCGTGTCAAGTCTCTTGGTGGCCCTCATC 622
 Db 301 AATTTCACAGTGGCCCGGTTTGGGGTAGCACGTTGCAAGTATCTTGGTGGCCCTCATC 360

QY 623 CAACCGCGCGGGCTGCATCTTCGCAGACAAGATCCATCTGCTTATGAGAGTGGCG 682
 Db 361 CAGCGCGTGGAGCTGCATCTTCGCAGACAAGATCCATCTGCTTATGAGAGTGGCGT 420

QY 683 TCTGGACCGTCTATCTTTAACTTCCCGGACCGGCAAGTGAAGTCAATCCCACTGTCTAC 742
 Db 421 TCTGGACCGTCTATCTTTAACTTCCCGGACCGGCAAGTGAAGTCAATCTGTCTAC 480

QY 743 CCGGGTGCAGTACATTTGTGCAATCATGATCGGCAATCTGAAAGGCACAAAATTTCTG 802
 Db 481 CCGGGTCCCGGACATTTGTGCAATCATGATTTGGCAATCTGAAAGGCACAAAATTTCTG 540

QY 803 CAATCTATTCAAGAGGCATCAAGTGACAATGTCATAGAGTAGGGAAGAAACATGGC 862
 Db 541 CATTTCTATTCAAGAGGCATCCAAGTCAATGTCATGGAAGTAGGGAAGAAACATGGC 600

QY 863 CTTGGGTGAATCACTATCAATTTTCTGTTCTGTCCTTTTATTATTATTCAGCGG 922
 Db 601 CTTGGGTGAATCACTATCAATTTTCTGTTCTGTCCTTTTATTATTATTCAGCGG 660

QY 923 GCAACTGGGGTATTTTATCTTTTATTTCTGTCGAGGCTACGGAATCAAGAGCTCAA 982
 Db 661 GCAACCGTAGCTATTTTATCTTTTATTTCTGTCGAGGATTCGCAATCAAGAGCTCAA 720

QY 983 AGCAGGAAGCAGAGGCAATTAAGGCAGATGCTTAAAGACTATTTGAAGGCTTCACTA 1042
 Db 721 AGGAGGAAGCAGAGGCAATTAAGGCAGATGCTTAAAGACTATTTGAAGGCTTCACTG 780

QY 1043 CGCACACTGAAACAAGGAGACAAGGAATTTGGCCCTGATGAGATAGTTGTGCTGTGTC 1102
 Db 781 CGCACCTTGAACAAGGAGACAAGGAATTTGGACCTGATGAGATAGTTGTGCTGTGTC 840

QY 1103 ATTGAATGTATAAACCAATGATTTGGTACGATCTTAAACGTCACCAATATTTTCCAT 1162
 Db 841 ATTGAGCTCTATAGCCAAATGATTTGGTGGCATCTTAACCTGAAATCATATTTTCCAT 900

QY 1163 AAGACATGCTGTGACCATGGCTTTAAACACAAGACTTGGCCCATGTGCAATGTGAC 1222
 Db 901 AATACATGTGTGGACCGCTGCTTTTAGAACAAGGACGTGGCCCATGTGCAAGTGAGAC 960

QY 1223 ATACTCAAGCTTTGGGAATTGAGGTGGATGTTGAAAGATGGATCAGTGTCTTTCAAGTC 1282
 Db 961 ATTCTCAAGCTCTTGGAAATTGAGGTGGATGTTGAAAGATGGATCAGTGTCTTTCAAGTT 1020

QY 1283 CCGTATATCAATGAAATATCTAATAGTGCCTCTCCCATGAAGAGGATAATTCGCGAGGAG 1342
 Db 1021 CCGTATATCAATGAAATATCTAATAGTGCCTCTCCCATGAAGAGGATAATTCGCGAGGAG 1080

QY 1343 ACCGATCATCTGATATGCTTACGTACAGGAAACAGATGAACCCCTCTTGGAGGAACAC 1402
 Db 1081 ACGGATCATCTGATATGATCATAGTCAAGGAGCAGATGAGCCACCTCTTGGAGGACCAT 1140

QY 1403 GTGCAGTCAACAAATGAAAGTCTACAGCTGTGTAACCATGAAGCAAAATTCGTGGCAGTG 1462
 Db 1141 GCGCAGTCAACAAAGGAAATCTACAGCTGTGTAACCATGAAGCAAAATTCGTGGCAGTG 1200

QY 1463 GATGTTATTCCTCATGTNGACCAACCCACTTTTGAAGAGAGCAAACTCTCTAATCAAGAG 1522
 Db 1201 GATGTTGTCCCCCATGTTGACCAACCCGACCTTTGAGAGAGATGAAATCTCTGATCAAGAG 1260

QY 1523 ACTGCTGTTCGAGAAATTAATCTTAA 1549
 Db 1261 GCTGCTGTTCGAGATTAATCTTAA 1287

RESULT 13

ABK12991
 ID ABK12991 standard; DNA; 1287 BP.

XX AC ABK12991;

XX DT 23-APR-2002 (first entry)

XX XX Synthetic goliath DNA sequence #5.

XX Goliath protein; antiangiogenic; vasotropic; gene therapy;
 dosage form; angiogenesis; neurogenesis; tumour; vascularisation;
 cancer; ischaemia; neuroblastoma; neurodegenerative disease;
 unrecovered nerve trauma; ds.

XX Synthetic.

XX WO200193681-A1.

XX PD 13-DEC-2001.

XX PF 01-JUN-2001; 2001WO-US18000.

XX PR 02-JUN-2000; 2000US-0586398.

XX PA (REGC) UNIV CALIFORNIA.

XX FI Harland R, Baker JC;

XX DR WPI; 2002-147637/19.

XX PT New compositions comprising goliath proteins, useful for modulating
 angiogenesis or neurogenesis in mammals e.g. for preventing or
 treating undesirable vascularisation of a tumour, ischaemia or
 neurodegenerative disease

XX PS Disclosure; Page 41-42; 45pp; English.

XX The present invention relates to a new pharmaceutical composition that
 comprises a goliath polypeptide in dosage form. The goliath polypeptide
 has a sequence identity of at least 75% to the protein sequences
 (AAU74918-AAU74921) fully defined in the specification. The composition
 is useful for modulating angiogenesis or neurogenesis in mammals,
 particularly in humans or mice. Specifically, the composition is useful
 for the prophylactic and/or therapeutic treatment of excess angiogenesis
 e.g. undesirable vascularisation of a tumour or insufficient angiogenesis
 e.g. ischaemia, excess neurogenesis e.g. neuroblastoma or insufficient
 neurogenesis e.g. neurodegenerative disease or unrecovered nerve trauma.

Db 686 GTACAGGGAACAGATGAACCGCCTCTGGAGGAACACACGTGCAGTCAACAAATGAAAGTCTA 745

The present invention relates to a new pharmaceutical composition that comprises a golliath polypeptide in dosage form. The golliath polypeptide has a sequence identity of at least 75% to the protein sequences (AAU74918-AAU74921) fully defined in the specification. The composition is useful for modulating angiogenesis or neurogenesis in mammals, particularly in humans or mice. Specifically, the composition is useful for the prophylactic and/or therapeutic treatment of excess angiogenesis e.g. undesirable vascularisation of a tumour or insufficient angiogenesis e.g. ischaemia, excess neurogenesis e.g. neuroblastoma or insufficient neurogenesis e.g. neurodegenerative disease or unrecovered nerve trauma.

CC The present nucleic acid sequence represents synthetic DNA sequence #3
CC that is one of several artificial goliath nucleic acids (ABK12987-
CC ABK12992) used in the invention for modulating angiogenesis or
CC neurogenesis.
XX

SQ Sequence 1287 BP; 308 A; 320 C; 355 G; 304 T; 0 other;

Query Match 56.3%; Score 999.6; DB 24; Length 1287;
Best Local Similarity 86.0%; Pred. No. 9.5e-268;
Matches 1107; Conservative 0; Mismatches 180; Indels 0; Gaps 0;

```
Qy 263 ATGGGGCGCGCGCTGCGGGCGGGGCTCTCTGCGCGGGTGGCTGCGGCTTTTCCAGATTG 322
Db 1 ATGGGGCGCGCGCGGGATGCGGCTCTAATGCGCGGGTGGCTGCGCGCTGCGCGGCTA 60
Qy 323 CTGCATGTTGCTTCTGCTGCGCTGAGTCCGACGACCCCGGTTCCCGGGGGGCTGAA 382
Db 61 CTGGCATGGTGTCTTCTTGCCCTGAGTCCGACGCGCAGGTTCCCGTGGAGCCGAC 120
Qy 383 GCAGTGTGGACCGCGTACCTCAACGTGTCTTGGCGGGTCCGCACACGGAGTGAACCGT 442
Db 121 GCCGTGTGGACCGCGTAACTCAACGTTTCTGGCGCGTTCGCGACACCGGAGTAAACCGC 180
Qy 443 ACGTGTGGAGCTGACGAGGAGGGCGGTGTACGGCCAGACCTCGCGCGCTGGAGCGTGTG 502
Db 181 ACTGTGTGGACCTGAGCGAGGAGGGCGGTATACGGCCATGACTCGCCCTGGAGCCGCTC 240
Qy 503 GCTGGGGTCTGTTGCTACCGCGCGACCGCGCGCGCGCTTAAACGCTGTAAACCGCACACG 562
Db 241 TCCGAGTCTGTTGCTGCGCGCGACGGCGCGCGCGCTCAAGCGCTGTATCCGACACC 300
Qy 563 AATTTCACGGTGCACCGTTTGGGGAAGCAACGTTGCAAGTCTCTTGGTTGGCCCTCATC 622
Db 301 AATTTCACGGTGCACAGTTTGGGGTAGCACCGTCCACGCTCAAGCTGTATCCGACATC 360
Qy 623 CAACGCGCGGGGCTGCACCTTCGCAGACAAAGATCCATCTGGCTTATGAGAGATGGGG 682
Db 361 CATCGGTGGCGGTGACGCTTCGCGAAGAATCCATCTGGCTTCGAGAGAGGGGCT 420
Qy 683 TCTGAGCGGTCACTCTTTAACTTCCCGGACCGCGAATGAGTCAATCCCGCATGTCTCAC 742
Db 421 TCTGAGCGGTCACTTTTAACTTCCCTGGGACCGCGAATGAAGTCACTCCCTATGTCTCAC 480
Qy 743 CCGGTGACGTAGACATTTGTCATCATGATCGGCAATCTGAAGGCCACAAAAATCTG 802
Db 481 CCGGTGCGGGGACATAGTTGCAATATGATTCGCAATCTGAAGGGAAACAAAAATCTG 540
Qy 803 CAATCTATTCAAAGAGGCATACAGTGAATGTCATAGAGTAGGGAAAAAACATGGC 862
Db 541 CATCTATTACAGAGGCATCGAAGTCAATGTCATTGAGTAGGCAAAAAACAGGGC 600
Qy 863 CTTGGGTGAATCACTATTCAATTTTTTTTTTCTGTGCTCTTTTATTTATTACGGCG 922
Db 601 CTTGAGTGAATCACTATTCAATCTTCTGCTGTGTCTGTCTCACTTTTCACTTTTACGGCC 660
Qy 923 GCAACTGTGGGCTATTTTATCTTTTATCTGCTCGAAGCTACGGAATGCAAGAGCTCAA 982
Db 661 GCAACCTGTGGGCTATTTATCTTTTATCTGCTCGAGATTACGGAATGCAAGAGCTCAA 720
Qy 983 AGCAGGAAGCAGAGGCAATTAAGGCAGATGCTAAAAAGCTATTGGAAGGCTTCAACTA 1042
Db 721 AGTAGGAAGCAGAGGCAATTGAAGGCAGAGCTAAAAATGCTATTGGCAAGCTTCAGCTG 780
Qy 1043 CGCAGCTGAACCAAGGAGACAGGAATTTGGCCCTGATGAGATAGTTGTGCTGTGTC 1102
Db 781 CGCAGCTGAACCAAGGAGACAGGAATTTGGCCCTGATGAGATAGCTGTGCTGTGTC 840
Qy 1103 ATTGAATTGTATAAACCAATGATTGTGTAGCATCTTAAACGTGCAACCATATTTTCCAT 1162
Db 841 ATTGAGCTGTATAAGCCAAATGATTTTGTGGCATCTCTAACTTGAATCATATATTCAT 900
Qy 1163 AAGACATGTGTGACCCCATGGCTGTATAAACCAAGACCTTGGCCCATGTGCAATGTGAC 1222
```

```
Db 901 AATACATGTGTCGACCGTGGCTTTTAGAAACACAGGACTTGCCTCATCTGAAAGTGGAC 960
Qy 1223 ATACTCAAAGCTTTTGGGAATTTGAGGTGATGTTGAAGATGGATCAGTGTCTTTTACAAGTC 1282
Db 961 ATTCCTAAAGCTCTTGGAAATTTGACGTGATGTTGAAGATGGATCAGTGTCTTTACAAGTC 1020
Qy 1283 CCTGTATCCAAATGAATATCTTAATAGTCCCTCTCCCATGAAGAGGATAATCGCAGCGAG 1342
Db 1021 CCTGTTTGAATGAAGCATCTAATACTGCTCTCTCCCATGAAGAGGACAGTCGAAAGTGAG 1080
Qy 1343 ACCGCATCATCTGCATATGCTTTCAGTACAGGACACAGATGAACCGCTCTGGAGGAACAC 1402
Db 1081 ACTGCATCATCCGATATGCGTCAAGGAGACAGATGAGCCACCCTCGGAGGACAT 1140
Qy 1403 GTGCAGTCAACAAATGAAGTCTACAGCTGGTAAACCATGAAGCAAAATCTGTGGCAGTG 1462
Db 1141 GCGCAATCAGCAATGAATCTCCAGTGTGAAACCATGAAGCAAAATCTGTGGCCGTC 1200
Qy 1463 GATGTTATTCCTCATGTNGACAAACCCAAACCTTTGAAAGAACGAACTCCTTAATCAAGAG 1522
Db 1201 GATGTTGTGCCCCATGTAGACAAACCTTACCTTTGACGAAGATGAGACTCCTGAAACAAG 1260
Qy 1523 ACTGCTGTTGAGAAATTTAAATCTTAA 1549
Db 1261 GCTGCTGTTGCGAGATTAAAGTCTTAA 1287
```

Search completed: April 26, 2003, 01:03:20
Job time : 299 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 26, 2003, 02:27:30 ; Search time 145 Seconds
(without alignments)
13312.692 Million cell updates/sec

Title: US-09-854-300-7

Perfect score: 1774

Sequence: 1 agctggagctccacgggt.....ataggcaagtctctctcag 1774

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 709820 seqs, 544064369 residues

Total number of hits satisfying chosen parameters: 1419640

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1178.8	66.4	2306	10	US-09-764-864-63
2	1157.6	65.3	1250	10	US-09-764-864-519
3	372.6	21.0	643	9	US-10-046-935-1815
4	372.6	21.0	643	9	US-09-878-178-1815
5	372.6	21.0	643	9	US-10-146-502-1815
6	310	17.5	435	10	US-09-920-300A-1176
7	310	17.5	435	12	US-10-033-528-1176
8	274.6	15.5	387	10	US-09-983-965-5551
9	224.6	12.7	616	9	US-10-060-038-25
10	214.4	12.1	374	10	US-09-983-965-4662
11	209	11.3	2063	10	US-09-935-390A-1
12	201.2	11.3	392	10	US-09-983-965-5467
13	187.4	10.6	340	10	US-09-783-590-3523
14	180.8	10.2	3615	9	US-09-822-846-596
15	149.2	8.4	1322	10	US-09-935-390A-16
16	132.6	7.5	619	10	US-09-764-864-485
17	132	7.4	508	9	US-09-918-995-21070
18	117.2	6.6	463	10	US-09-864-761-11364
19	112.8	6.4	826	10	US-09-822-849A-90

ALIGNMENTS

RESULT 1

US-09-764-864-63

; Sequence 63, Application US/09764864

; Patent No. US20020132753A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PTZ23

; CURRENT APPLICATION NUMBER: US/09/764,864

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - consult PALM or file wrapper

; NUMBER OF SEQ ID NOS: 1792

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 63

; LENGTH: 2306

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-764-864-63

Query Match 66.4%; Score 1178.8; DB 10; Length 2306;

Best Local Similarity 99.0%; Pred. No. 4.1e-311;

Matches 1217; Conservative 0; Mismatches 8; Indels 4; Gaps 3;

Qy	547	CTGTAAACCGCACACGAATTTTCACGGTGCCACGGTTTGGGGAAGCACCGTGCAAGTCTC	606
Db	12	CTGTAAACCGCACACGAATTTTCACGGTGCCACGGTTTGGGGAAGCACCGTGCAAGTCTC	71
Qy	607	TTGGTTGGCCCTCATCAACGCGCGGGGGCTGCACCTTCGACACAAGATCCATCTGGC	666
Db	72	TTGGTTGGCCCTCATCAACGCGCGGGGGCTGCACCTTCGACACAAGATCCATCTGGC	131
Qy	667	TTATGAGATGGGGCTGTGGAGCCGTCATCTTTAACTTCCCGGGACCCGCAATGAGGT	726
Db	132	TTATGAGATGGGGCTGTGGAGCCGTCATCTTTAACTTCCCGGGACCCGCAATGAGGT	191
Qy	727	CATCCCCATGCTCACCCGGGTGCAGTAGACATTTTGGCAATCATGATCGGCAATCTGAA	786
Db	192	CATCCCCATGCTCACCCGGGTGCAGTAGACATTTTGGCAATCATGATCGGCAATCTGAA	251
Qy	787	AGGCACAAAATCTTCAATCTATTCAAGAGGCATACAAATGATGATAGTATAGAGT	846
Db	252	AGGCACAAAATCTTCAATCTATTCAAGAGGCATACAAATGATGATAGTATAGAGT	311

Sequence 5079, Ap
Sequence 1328, Ap
Sequence 27944, A
Sequence 240, App
Sequence 23, Appl
Sequence 7, Appl
Sequence 484, App
Sequence 36, Appl
Sequence 79, Appl
Sequence 81, Appl
Sequence 222, App
Sequence 222, App
Sequence 344, App
Sequence 182, App
Sequence 182, App
Sequence 404, App
Sequence 466, App
Sequence 40, Appl
Sequence 148, App
Sequence 16, Appl
Sequence 686, App
Sequence 231, App
Sequence 32, Appl
Sequence 32, Appl
Sequence 103, App

QY 847 AGGAAAAACATGGCCCTTGGGTGAATCACTATTCAA-TTTTTTTCGTTCTGTGCTCT 905
Db 312 AGGAAAAACATGGCCCTTGGGTGAATCACTATTCAA-TTTTTTTCGTTCTGTGCTCT 371
QY 906 TTTTATTTATTAACGGGGCAACTGTGGCTATTTATCTTTTATTTCTGTCTGAAGCTAC 965
Db 372 TTTTATTTATTAACGGGGCAACTGTGGCTATTTATCTTTTATTTCTGTCTGAAGCTAC 431
QY 966 GGAATGCAAGAGCTCAAGCAGGAGCAGGCAATTAAGGCAGATGCTTAAAGAGCTA 1025
Db 432 GGAATGCAAGAGCTCAAGCAGGAGCAGGCAATTAAGGCAGATGCTTAAAGAGCTA 491
QY 1026 TTGGAAGCTTTCAACTACGCACTGTAAGCAAGGAGCAGGAAATTTGGCCCTGTATGGAG 1085
Db 492 TTGGAAGCTTTCAACTACGCACTGTAAGCAAGGAGCAGGAAATTTGGCCCTGTATGGAG 551
QY 1086 ATAGTTGTGCTGTGCTGAATTTGATTAACCAATGATTTGGTACGCATCTTAACGT 1145
Db 552 ATAGTTGTGCTGTGCTGAATTTGATTAACCAATGATTTGGTACGCATCTTAACGT 611
QY 1146 GCAACCATATTTTCCATAAGACATGTTTGACCCATGGCTTTAAACCAAGACTTTGCC 1205
Db 612 GCAACCATATTTTCCATAAGACATGTTTGACCCATGGCTTTAAACCAAGACTTTGCC 671
QY 1206 CCATGTGCAAAATGTGACATCTCAAGCTTTGGGAATTTGAGTGGATTTGAAGATGGAT 1265
Db 672 CCATGTGCAAAATGTGACATCTCAAGCTTTGGGAATTTGAGTGGATTTGAAGATGGAT 731
QY 1266 CAGTGTCTTTACAGTCCCTGTATCCAAATGAAATATCTAATAGTGGCTCTCCCATGAAG 1325
Db 732 CAGTGTCTTTACAGTCCCTGTATCCAAATGAAATATCTAATAGTGGCTCTCCCATGAAG 791
QY 1326 AGGATAATTCGAGGAGACCGCATCTCTGATATGCTTCAAGTACAGGGAACAGATGAAC 1385
Db 792 AGGATAATTCGAGGAGACCGCATCTCTGATATGCTTCAAGTACAGGGAACAGATGAAC 851
QY 1386 CGCTCTGAGGAAACAGCTGAGTCAACAAATGAAAGTCTACAGTGGTAAACCATGAAG 1445
Db 852 CGCTCTGAGGAAACAGCTGAGTCAACAAATGAAAGTCTACAGTGGTAAACCATGAAG 911
QY 1446 CAATCTGTGGAGTGGATTTATCTCTCATGTGACACCAACCTTTGAGAGAGAG 1505
Db 912 CAATCTGTGGAGTGGATTTATCTCTCATGTGACACCAACCTTTGAGAGAGAG 971
QY 1506 AAATCTCTTAATCAAGAGACTGCTGTTCGAGAAATTAATCTTAAATCTGTGTAATAGA 1565
Db 972 AAATCTCTTAATCAAGAGACTGCTGTTCGAGAAATTAATCTTAAATCTGTGTAATAGA 1031
QY 1566 AAATCTCTTAATCAAGAGACTGCTGTTCGAGAAATTAATCTTAAATCTGTGTAATAGA 1625
Db 1032 AAATCTCTTAATCAAGAGACTGCTGTTCGAGAAATTAATCTTAAATCTGTGTAATAGA 1091
QY 1626 TGGATAAATTTAATAAATAAGAGTGTATCTGAAAGTCTCAGATGACTAATATTATGCT 1685
Db 1092 TGGATAAATTTAATAAATAAGAGTGTATCTGAAAGTCTCAGATGACTAATATTATGCT 1151
QY 1686 ATAGTTAAATGGCTTAAATAATTTAACTGTTAACTTTTTCACCAAACTCATATA 1745
Db 1152 ATAGTTAAAT--TGGCTTAAATAATTTAACTGTTAACTTTTTCCTCA--CAAACTCATATA 1208
QY 1746 ATATTTTTCATAGGCAAGTTTCTCTCAG 1774
Db 1209 ATATTTTTCATAGGCAAGTTTCTCTCAG 1237

RESULT 2

US-09-764-864-519

; Sequence 519, Application US/09764864

; Patent No. US20020132753A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PT223
CURRENT APPLICATION NUMBER: US/09/764,864
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 1792
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 519
LENGTH: 1250

TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (540)
OTHER INFORMATION: n equals a.t.g. or c
NAME/KEY: SITE
LOCATION: (1242)
OTHER INFORMATION: n equals a.t.g. or c
US-09-764-864-519

Query Match 65.3%; Score 1157.6; DB 10; Length 1250;
Best Local Similarity 98.8%; Pred. No. 1.7e-305;
Matches 1208; Conservative 0; Mismatches 10; Indels 5; Gaps 4;

QY 547 CTGTAAACCGCAGACAGAAATTTACGGTCCCGGCTGGGGAGGACCGTGCAGCTC 606
Db 18 CTGTAAACCGCAGACAGAAATTTACGGTCCCGGCTGGGGAGGACCGTGCAGCTC 77
QY 607 TTGGTTGGCCCTCATCAACGCGCGGGGCTGCACCTTCGACAGCAAGATCCATCTGGC 666
Db 78 TTGGTTGGCCCTCATCAACGCGCGGGGCTGCACCTTCGACAGCAAGATCCATCTGGC 137
QY 667 TTATGAGAGATGGCGCTCTGGAGCGCTCATCTTTAACTTCCCGGAGCCGCAATAGGT 726
Db 138 TTATGAGAGAGGGCGCTCTGGAGCGCTCATCTTTAACTTCCCGGAGCCGCAATAGGT 197
QY 727 CATCCCATGCTCTCACCGGGTGCAGTAGACATTTGTCGAATCATGATCGCAATCTGAA 786
Db 198 CATCCCATGCTCTCACCGGGTGCAGTAGACATTTGTCGAATCATGATCGCAATCTGAA 257
QY 787 AGGCACAAAAATTTCTGCAATCTATTCAAGAGGCATACAAAGTGAATGGTCTAGAGT 846
Db 258 AGGCACAAAAATTTCTGCAATCTATTCAAGAGGCATACAAAGTGAATGGTCTAGAGT 317
QY 847 AGGAAAAACATGGCCCTTGGGTGAATCACTATTCAA-TTTTTTTCGTTCTGTGCTCT 905
Db 318 AGGAAAAACATGGCCCTTGGGTGAATCACTATTCAA-TTTTTTTCGTTCTGTGCTCT 377
QY 906 TTTTATTTATTAACGGGGCAACTGTGGCTATTTATCTTTTATTTCTGCTCGAAGGCTAC 965
Db 378 TTTTATTTATTAACGGGGCAACTGTGGCTATTTATCTTTTATTTCTGCTCGAAGGCTAC 437
QY 966 GGAATGCAAGAGCTCAAGCAGGAGCAGGCAATTAAGGCAGATGCTTAAAGAGCTA 1025
Db 438 GGAATGCAAGAGCTCAAGCAGGAGCAGGCAATTAAGGCAGATGCTTAAAGAGCTA 497
QY 1026 TTGGAAGCTTTCAACTACGCACTGTAAGCAAGGAGCAGGAAATTTGGCCCTGTATGGAG 1084
Db 498 TTGGAAGCTTTCAACTACGCACTGTAAGCAAGGAGCAGGAAATTTGGCCCTGTATGGAG 557
QY 1085 GATAGTTGTGCTGTGCTGAATTTGATTAACCAATGATTTGGTACGCATCTTAACG 1144
Db 558 GATAGTTGTGCTGTGCTGAATTTGATTAACCAATGATTTGGTACGCATCTTAACG 617
QY 1145 TGCAACCATATTTTCCATAAGACATGTTTGACCCATGGCTTTAAACACAGACTTGC 1204
Db 618 TGCAACCATATTTTCCATAAGACATGTTTGACCCATGGCTTTAAACACAGACTTGC 677
QY 1205 CCATGTGCAAAATGTGACATCTCAAGCTTTGGGAATTTGAGGTGGATGTTGAAGATGGA 1264
Db 678 CCATGTGCAAAATGTGACATCTCAAGCTTTGGGAATTTGAGGTGGATGTTGAAGATGGA 737
QY 1265 TCAGTGTCTTTACAGTCCCTGTATCCCAATGAAATATCTAATAGTGGCTCTCCCATGAA 1324

Db 738 TCAGTGTCTTTTACAGTCCCTGTATCAATGAATATCTAATAGTCCCTCCCTCCATGAA 797
Qy 1325 GAGGATATCCGACGACCGCATCATCTGGATATCTTCAGTACAGGGAACAGATGAA 1384
Db 798 GAGGATATCCGACGACCGCATCATCTGGATATCTTCAGTACAGGGAACAGATGAA 857
Qy 1385 CCGCTCTGGAGGAACACGTCAGTCAACAAATGAAGTCTACAGCTGGTAAACCATGAA 1444
Db 858 CCGCTCTGGAGGAACACGTCAGTCAACAAATGAAGTCTACAGCTGGTAAACCATGAA 917
Qy 1445 GCAATCTGCGGAGTGGATGTTATCTCTCATGTGACACCAACCTTTGAGGAGAC 1504
Db 918 GCAATCTGCGGAGTGGATGTTATCTCTCATGTGACACCAACCTTTGAGGAGAC 977
Qy 1505 GAAACTCTCTAATCAAGAGACTGCTGTTTCGAGAAATTAATCTTAAATCTGTGTAATAG 1564
Db 978 GAAACTCTCTAATCAAGAGACTGCTGTTTCGAGAAATTAATCTTAAATCTGTGTAATAG 1037
Qy 1565 AAACTTGAACCATAGTAATAACAGAACTGCCAATCAGGCGCTAGTTCTTATTAATAAA 1624
Db 1038 AAACTTGAACCATAGTAATAACAGAACTGCCAATCAGGCGCTAGTTCTTATTAATAAA 1097
Qy 1625 TTGATATAATTTAAATAAAGAGTGATCTGAAAGTGTCTCAGATGACTAATATTATGC 1684
Db 1098 TTGATATAATTTAAATAAAGAGTGATCTGAAAGTGTCTCAGATGACTAATATTATGC 1157
Qy 1685 TATAGTTAAATGGCTTAAATAATTTAACTGTAACTTTTTTCCACCAACTCAATTAT 1744
Db 1158 TATAGTTAAA--TGGCTTAAATAATTTAACTGTAACTTTTTTCCCA-CAAACTCAATTAT 1214
Qy 1745 AATATTTTTCATAGCAAGTTTC 1767
Db 1215 AATATTTTTCATAGCAATTTCC 1237

RESULT 3

US-10-046-935-1815

; Sequence 1815, Application US/10046935
; Patent No. US2002015601A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secretist, Heather
; APPLICANT: Wang, Aijun A.
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.527C1
; CURRENT APPLICATION NUMBER: US/10/046.935
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 2239
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1815
; LENGTH: 643
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 584
; OTHER INFORMATION: n = A,T,C or G
US-10-046-935-1815

Query Match 21.0%; Score 372.6; DB 9; Length 643;
Best Local Similarity 98.0%; Pred. No. 2.1e-91;
Matches 398; Conservative 0; Mismatches 5; Indels 3; Gaps 2;

Qy 1369 ACAGGGAACAGATGAACCGCTCTCGAGGAACACGTCGAGTCAACAAATGAAGTCTACA 1428
Db 1 ACAGGGAACAGATGAACCGCTCTCGAGGAACACGTCGAGTCAACAAATGAAGTCTACA 60
Qy 1429 GCTGTGTAACCATGAAGCAAAATCTGTGGCAGTGGATGTTATCTCATGTGNGACAACC 1488
Db 61 GCTGTGTAACCATGAAGCAAAATCTGTGGCAGTGGATGTTATCTCATGTGNGACAACC 120

Qy 1489 AACCTTTGAAGAACGAAACTCTTAATCAAGAGACTGCTCTCGAGAAATTAATCTTA 1548
Db 121 AACCTTTGAAGAACGAAACTCTTAATCAAGAGACTGCTCTCGAGAAATTAATCTTA 180
Qy 1549 AAATCTGTGTAATGAAGAACTTGAACCATTAATGAAGTCTGCAATCAGGGCT 1608
Db 181 AAATCTGTGTAATGAAGAACTTGAACCATTAATGAAGTCTGCAATCAGGGCT 240
Qy 1609 AGTTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1668
Db 241 AGTTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 300
Qy 1669 ATGACTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1728
Db 301 ATGACTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 358
Qy 1729 CCACCAAACTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1774
Db 359 CCA-CAAACTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 403

RESULT 4

US-09-878-178-1815

; Sequence 1815, Application US/09878178
; Patent No. US2002017552A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secretist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.527
; CURRENT APPLICATION NUMBER: US/09/878.178
; CURRENT FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 2237
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1815
; LENGTH: 643
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)-(643)
; OTHER INFORMATION: n = A,T,C or G
US-09-878-178-1815

Query Match 21.0%; Score 372.6; DB 9; Length 643;
Best Local Similarity 98.0%; Pred. No. 2.1e-91;
Matches 398; Conservative 0; Mismatches 5; Indels 3; Gaps 2;

Qy 1369 ACAGGGAACAGATGAACCGCTCTCGAGGAACACGTCGAGTCAACAAATGAAGTCTACA 1428
Db 1 ACAGGGAACAGATGAACCGCTCTCGAGGAACACGTCGAGTCAACAAATGAAGTCTACA 60
Qy 1429 GCTGTGTAACCATGAAGCAAAATCTGTGGCAGTGGATGTTATCTCATGTGNGACAACC 1488
Db 61 GCTGTGTAACCATGAAGCAAAATCTGTGGCAGTGGATGTTATCTCATGTGNGACAACC 120
Qy 1489 AACCTTTGAAGAACGAAACTCTTAATCAAGAGACTGCTCTCGAGAAATTAATCTTA 1548
Db 121 AACCTTTGAAGAACGAAACTCTTAATCAAGAGACTGCTCTCGAGAAATTAATCTTA 180
Qy 1549 AAATCTGTGTAATGAAGAACTTGAACCATTAATGAAGTCTGCAATCAGGGCT 1608
Db 181 AAATCTGTGTAATGAAGAACTTGAACCATTAATGAAGTCTGCAATCAGGGCT 240
Qy 1609 AGTTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1668
Db 241 AGTTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 300
Qy 1669 ATGACTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1728
Db 359 CCA-CAAACTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 403

Db 301 ATGACTAATATATGCTATAGTAAA--TGGCTTAAATATTTAACTGTTAACTTTTT 358
QY 1729 CCACAAACTATTAATATATTTTCATAGGCAAGTTTCCTCTCAG 1774
Db 359 CCA-CAAACTATTATAATTTTTCATAGGCAAGTTTCCTCTCAG 403

RESULT 5

US-10-146-502-1815
; Sequence 1815, Application US/10146502
; Publication No. US20030069180A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Harlocker, Susan M.
; APPLICANT: Secrist, Heather
; APPLICANT: Wang, Aijun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.527C2
; CURRENT APPLICATION NUMBER: US/10/146,502
; CURRENT FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1815
; LENGTH: 643
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 594
; OTHER INFORMATION: n = A,T,C or G
US-10-146-502-1815

Query Match 21.0%; Score 372.6; DB 9; Length 643;
Best Local Similarity 98.0%; Pred. No. 2.1e-91;
Matches 398; Conservative 0; Mismatches 5; Indels 3; Gaps 2;

QY 1369 ACAGGGAACAGATGAACCGCCTCTGGAGGAACACGTCGACGCAACCAATGAAGTCTACA 1428
Db 1 ACAGGGAACAGATGAACCGCCTCTGGAGGAACACGTCGACGCAACCAATGAAGTCTACA 60
QY 1429 GCTGTAAACCATGAAGCAATCTGTGGCAGTGGATGTTATTCCTCATGTNGACAACC 1488
Db 61 GCTGTAAACCATGAAGCAATCTGTGGCAGTGGATGTTATTCCTCATGTNGACAACC 120
QY 1489 AACCTTTGAAGAGACGAAACTCCTAATCAAGAGACTGCTGTTTCGAGAAATTTAAATCTTA 1548
Db 121 AACCTTTGAAGAGACGAAACTCCTAATCAAGAGACTGCTGTTTCGAGAAATTTAAATCTTA 180
QY 1549 AAATCTGTGTAATAGAAACTTTGAACCATTAGTAATACAGAACTGCCAATCAGGGCCT 1608
Db 181 AAATCTGTGTAATAGAAACTTTGAACCATTAGTAATACAGAACTGCCAATCAGGGCCT 240
QY 1609 AGTTTCTATTAATAATTTGGATAAATTTAATAAATAAGAGTGATCTCAAAAGTGTCTAG 1668
Db 241 AGTTTCTATTAATAATTTGGATAAATTTAATAAATAAGAGTGATCTCAAAAGTGTCTAG 300
QY 1669 ATGACTAATATATGCTATAGTTAAATGGCTTTAAAAATTAATTAACCTGTTAACTTTTT 1728
Db 301 ATGACTAATATATGCTATAGTTAAATGGCTTTAAAAATTAATTAACCTGTTAACTTTTT 358
QY 1729 CCACCAACTCATTATAATATTTTTCATAGGCAAGTTTCCTCTCAG 1774
Db 359 CCA-CAAACTATTATAATTTTTCATAGGCAAGTTTCCTCTCAG 403

RESULT 6

US-09-920-300A-1176/c
; Sequence 1176, Application US/09920300A
; Patent No. US20020136728A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.

; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547
; CURRENT APPLICATION NUMBER: US/09/920,300A
; CURRENT FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 1789
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1176
; LENGTH: 435
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-920-300A-1176

Query Match 17.5%; Score 310; DB 10; Length 435;
Best Local Similarity 97.4%; Pred. No. 2e-74;
Matches 336; Conservative 0; Mismatches 6; Indels 3; Gaps 2;

QY 1430 CTGGTAAACCATGAAGCAAAATCTGTGGCAGTGGATGTTATTCCTCATGTNGACAACCA 1489
Db 435 CTGGTAAACCATGAAGCAAAATCTGTGGCAGTGGATGTTATTCCTCATGTNGACAACCA 376
QY 1490 ACCTTTGAAGAGACGAAACTCCTAATCAAGAGACTGCTGTTTCGAGAAATTTAAATCTTAA 1549
Db 375 ACCTTTGAAGAGACGAAACTCCTAATCAAGAGACTGCTGTTTCGAGAAATTTAAATCTTAA 316
QY 1550 AATCTGTGTAATAGAAACTTTGAACCATTTAGTAATAACAGAACTGCCAATCAGGGCCTA 1609
Db 315 AATCTGTGTAATAGAAACTTTGAACCATTTAGTAATAACAGAACTGCCAATCAGGGCCTA 256
QY 1610 GTTCTTATTAATAATTTGGATAAATTTAATAAATAAGAGTGATCTGAAAGTGTCTCAGA 1669
Db 255 GTTCTTATTAATAATTTGGATAAATTTAATAAATAAGAGTGATCTGAAAGTGTCTCAGA 196
QY 1670 TGACTAATATATGCTATAGTTAAATGGCTTAAATATTTAAATATTTAACTGTTAACTTTTTTC 1729
Db 195 TGACTAATATATGCTATAGTTAAAT--TGGCTTAAATATTTAACTGTTAACTTTTTTC 138
QY 1730 CACCAAACTCATTATATATTTTTCATAGGCAAGTTTCCTCTCAG 1774
Db 137 CA-CAAACTCATTATATATTTTTCATAGGCAAGTTTCCTCTCAG 94

RESULT 7

US-10-033-528-1176/c
; Sequence 1176, Application US/10033528
; Patent No. US20020131971A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.

; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547C1
; CURRENT APPLICATION NUMBER: US/10/033,528
; CURRENT FILING DATE: 2001-12-26
; NUMBER OF SEQ ID NOS: 1896
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1176
; LENGTH: 435
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-033-528-1176

Query Match 17.5%; Score 310; DB 12; Length 435;
Best Local Similarity 97.4%; Pred. No. 2e-74;
Matches 336; Conservative 0; Mismatches 6; Indels 3; Gaps 2;

QY 1430 CTGGTAAACCATGAAGCAAAATCTGTGGCAGTGGATGTTATTCCTCATGTNGACAACCA 1489
Db 435 CTGGTAAACCATGAAGCAAAATCTGTGGCAGTGGATGTTATTCCTCATGTNGACAACCA 376

; APPLICANT: Byatt, John C.
 ; APPLICANT: Mathialagan, Nagappan
 ; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
 ; FILE OF INVENTION: MUSCLE AND FAT DEPOSITION
 ; FILE REFERENCE: 37-21(10297)C
 ; CURRENT APPLICATION NUMBER: US/09/983,965
 ; CURRENT FILING DATE: 2001-10-26
 ; PRIOR APPLICATION NUMBER: US 09/465,231
 ; PRIOR FILING DATE: 1999-12-15
 ; PRIOR APPLICATION NUMBER: US 60/113,678
 ; PRIOR FILING DATE: 1998-12-17
 ; NUMBER OF SEQ ID NOS: 5912
 ; SEQ ID NO 4662
 ; LENGTH: 374
 ; TYPE: DNA
 ; ORGANISM: Bos taurus
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 14-LIB34-009-Q1-E1-D5
 US-09-983-965-4662

Query Match 12.1%; Score 214.4; DB 10; Length 374;
 Best Local Similarity 78.2%; Pred. No. 2.3e-48;
 Matches 283; Conservative 0; Mismatches 76; Indels 3; Gaps 2;
 QY 75 CGAGGAGTGCATCTCGCGCAACCTGTGTGTGCTGACGTACGTGCTCTCTGCTCGGAGCT 134
 Db 13 CAAGGGCTTCTCTGGCGACACATCTCGGCGACGCTACGTGCCCGCGGCTCTGACCC 72
 QY 135 AGCTGCGAGCTCC--CCAGTCTCACTCAATTCCTTCCCGACCTGGCGGCGACCTGCTCAA 192
 Db 73 AGCCCGCCCGCGCGCGAGTCTCGTCCGCTCTCTCACTACCTGGCGGCGACCTGGCCCA 132
 QY 193 GACGAGGCTCTGCGCAAGCGCTAGGAGGCGCGTGCACAGGCGCTAGCGAACTGCGGAG 252
 Db 133 GCGCGGGTCCCGCGAGAGTCCGAGAGCGGCTGCGAGGGCGCTTAGGAACTCAGAG 192
 QY 253 CGCGCGCGCATGCGGCGCGCGCTGCGGCGCGGCTGCTGCTGCGGCGGCTGCTGCGGCTT 312
 Db 193 CGCGCATATCATGCGGCGAGCTGCCAGGGCGCGGCTTCTGCGCGGCTGCTGCGGCTT 252
 QY 313 TTCCAGATTGCTGGCATGTGCTTCTGCTGCGCGCTGAGTCCGAGGCGACCGGTTCCCG 372
 Db 253 TTCCCGATTGCTGGCTGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 312
 QY 373 GCGGCGTGAAGCAGTGTGACCGCGCTACCTCAACGT-PTCTGCGCGGTTCCGACACCG 431
 Db 313 TGGAGCGGAGCGGTGTGGAGCGGTAGCTCAACGTGTCTTCCGCGGCTTCCGACACCG 372
 QY 432 GA 433
 Db 373 GA 374

RESULT 11

US-09-935-390A-1
 ; Sequence 1, Application US/09935390A
 ; Patent No. US20020076761A1

GENERAL INFORMATION:

; APPLICANT: Escobedo, Jaime
 ; Quianjin, Hu
 ; Garcia, Pablo
 ; Williams, Lewis T.

; TITLE OF INVENTION: Secreted Human Proteins

; NUMBER OF SEQUENCES: 38

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Chiron Corporation

; STREET: 4560 Horton Street

; CITY: Emeryville

; STATE: CA

; COUNTRY: USA

; ZIP: 94608-2916

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/935,390A
 ; FILING DATE: 22-Aug-2001
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/988,671
 ; FILING DATE: 1997-12-11
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Jane E. R. Potter
 ; REGISTRATION NUMBER: 33,332
 ; REFERENCE/DOCKET NUMBER: 1369.002
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (510) 923-2718
 ; TELEFAX: (510) 655-3542
 ; TELETYPE: <Unknown>
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2063 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; FEATURE:
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-09-935-390A-1
 Query Match 11.8%; Score 209; DB 10; Length 2063;
 Best Local Similarity 54.6%; Pred. No. 2.1e-46;
 Matches 507; Conservative 0; Mismatches 410; Indels 12; Gaps 4;
 QY 339 TGCTGGCCCTGAGTCCGCGAGCACCCGCTTCCCGGGGGCTGAAAGCAGTGTGACCGCGT 398
 Db 180 TGCTGGCCCTGCGCCCTGTGCTGCGCGCGCGCGCGCGCGCTCTCGAGTGTCTCGG 239
 QY 399 ACCTCAAGCTGTCTCTGGCGGGTTCGCCACACGGGAGTGAACCGTAGCGTGTGGAGCTGA 458
 Db 240 CCGT---GGTAAACATCGAGTAGCTGGACCCCGCAGACCAACCTGACGGTGTGGAGCT 296
 QY 459 GCAGGAGGCGGTGTACGGCCAGACTCGCGCGCTGAGAGCTGTGGCTGGGGTCTCTGGTAC 518
 Db 297 CGAGAGTGGCCCTTGGCGGACAGCTCGCCCAAGAGGGCGGCGCATGCGCTGTGGCGG 356
 QY 519 CGCCCGACGGCGCGCGGGCGCTTAAACGCTGTAAACCGCACACGAAATTTACCGTGCCCA 578
 Db 357 TCCCGTGGCGCGCGCGGAGACCTCGAGGGCTGCGCGCGCGCGCGCGCTTCTTCTGCTGC 416
 QY 579 CGGTTTGGGGAAGCACCGTCAAGTCTCTTGGTTGGCCCTCATCCACCGCGCGGGGCT 638
 Db 417 CCGAGCGCGCGCGCGGCGCGCGCTGGGCTGGGCTGCGCTCTGTTGGCTC---GTGGGGGCT 473
 QY 639 GCACCTTTCGCAGACAGATCCATCTGGCTTATGAGAGATGGGCTCTGGAGCGCTCATCT 698
 Db 474 GCACCTTCAAGGACAGGTGCTGTGCGCGCGCGGAGAACGCTCTGGCGCTGCTCTCT 533
 QY 699 TTAACCTTCCCGGACCGGCAATGAGTCAATCCCATGTCTCACCGGGTGGAGTAGACA 758
 Db 534 ACAATGAGGAGCGCTACGGGAACATCACCTTGGCCCATGTCTCACCGGGGAACAGGAAATA 593
 QY 759 TTGTTGCAATCATGATCGGCAATCTGAAAGGCACAAAAATTTGCAATCTATTCAAAGAG 818
 Db 594 TAGTGGTCAATATGATTAGCTATCCAAAGAGAGAGAAATTTTGGAGCTGGTGCAAAAG 653
 QY 819 GCATACAGTGACAATGGTCAAGAGTATAGAGTATGAGGAAAAAATGCGCCCTTGGGTGAATCAC- 877
 Db 654 GAATTCAGTAAACGATGACCATAGGGGTGGGCAACCGGCGCATGTACAGAGTGTTCATCAGC 713
 QY 878 --TATTCAATTTTTCGTTTCTGTGCTCTTTTATTATTATAGCGCGCAACTGTGGCT 935
 Db 714 GTCAGTCTGTGGTGTGTTGGGCCAATGCTTTCATCCCATGATTAATCTCTGTAGCCT 773

QY 936 ATTTTATCTTTTATCTGCTGAGGCTACGGAATGCAAGAGCTCAAGAGGAGGAGCAGA 995
Db 774 GGCTAATATTTTACTATATACAGCGTTTCTATATATCTGCTCAGATTTGGAAGTCA 833
QY 996 GCGAATTAAGGCGAGATGCTTAAAGAGCTATTGGAAGGCTTCAACTACGACACTGAAC 1055
Db 834 GCGATAGAA--AGAACTAGAAAGTTATTGGCCAGCTTCTACTTCATCTGTAAGC 890
QY 1056 AAGGAGACAGGAAATGGCCCTGATGGAGATAGTTGTCTGTGTCATTTGAATTTGATA 1115
Db 891 ATGGAGAAAGGAATGATGTTGATGCTGAAATTTGTGCTGCTGATTTGAATTTCA 950
QY 1116 AACCAATGATTTGGTACGCTCTTAAGCTGCAACCATTTTTCATAGACATGTTG 1175
Db 951 AAGTAAGGATATTAATAGAAATTTGCCATGCAAGCATATTTTTCATAGAATATGATTG 1010
QY 1176 ACCATGGCTGTTAAACACAGACTTGGCCCATGTGCCATGTGCAATGTCATCTCAAGCTT 1235
Db 1011 ACCATGGCTTTTGGATCAGGACATGTCATGTTGAATTTGATGTCATCAAGGCC 1070
QY 1236 TGGGAATGAGTGGGATGTTGAAGATGGA 1264
Db 1071 TAGGATATTGGGAGAGAGCTGGGATGTA 1099

RESULT 12

US-09-983-965-5467
; Sequence 5467, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Rao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 37-21(10297)C
; CURRENT APPLICATION NUMBER: US/09/983,965
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 09/465,231
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/113,678
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 5467
; LENGTH: 392
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; OTHER INFORMATION: Clone ID: 46-LIB34-036-Q1-E1-D10
US-09-983-965-5467

Query Match 11.3%; Score 201.2; DB 10; Length 392;
Best Local Similarity 80.3%; Pred. No. 9.3e-45;
Matches 285; Conservative 0; Mismatches 64; Indels 6; Gaps 4;

QY 1418 GAAAGTCTACAGCTGTTAAACCATGAAGCAATTTCTGTCGACGTGATGTTATCTCTAT 1477
Db 1 GAAATCTCAGCTGTTAAACCATGAAGCAATTTCTGTCGACGTGATGTTATCTCTAC 60
QY 1478 GTNGACAACCCAACTTTGAAGAGAGCAAACTCTTAATCAAGAGACTGCTGTTGAGAA 1537
Db 61 GTTGACAACCCAACTTTGAAGAGAGATGAAGTCTCTGATCAGGAAC-AATGTCAGAA 119
QY 1538 ATTAATCTTAAATCTGTGTAATAGAAAACTTGAACCAATAGTAATTAACGAAGCTGCC 1597
Db 120 ATTAATCTTAAATCTGTGTAATAGAAAACTTGAACCAATAGTAATTAACGAAGCTGCC 179
QY 1598 AATC-AGGGCTAGTTCTTATTAATAAATTTGGATAAATTTAATAAATAGAGTATCT 1656
Db 180 AATCAAGGGCCAATTTACAAACAAGGGAATGGGTAACCTTTAATAATAGAAAGACT 239
QY 1657 GAAAGTCTCAGATGACTAATATTTATGCTATAGTTAAAAATGGCTTAAAAATTTAAACCT 1716

Db 240 GAAATGCTGATATGACTAATATTTATGCTATAGTTAAA--TGACTTAAAAATCTAAACCT 297
QY 1717 GTTAACTTTTCCACCAACTCATTAATATTTTTCATAGGCAAGTTTCTCT 1771
Db 298 ATTAACCTTTT--CCACAAACACATTATATAAGTTTTTTCATAGGCAAAATTTCTCT 350

RESULT 13

US-09-783-590-3523
; Sequence 3523, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3523
; LENGTH: 340
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (205)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (220)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (221)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (244)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (265)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (294)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (318)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (326)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (332)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (333)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-3523

Query Match 10.6%; Score 187.4; DB 10; Length 340;
Best Local Similarity 97.0%; Pred. No. 5e-41;
Matches 191; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1203 GCCCATGTGCAATGTGACATCTCAAGCTTTGGGAATTTGAGTGGATGTTGAAGATG 1262
Db 1 GGACAGGGGCAATGTGACATCTCAAGCTTTGGGAATTTGAGTGGATGTTGAAGATG 60

Qy	1263	GATCAGTGCTTTTACAAGTCCCTGTCATCAATGAAATATCTTAATAGTGCCTCTCCCGATG	1322
Db	61	GATCAGTGCTTTTACAAGTCCCTGTCATCAATGAAATATCTTAATAGTGCCTCTCCCGATG	120
Qy	1323	AAGAGGATAATCGCAGCGAGACCGCATCATCTGGATATGCTTTCAGTACGAGAACAGATG	1382
Db	121	AAGAGGATAATCGCAGCGAGACCGCATCATCTGGATATGCTTTCAGTACGAGAACAGATG	180
Qy	1383	AACCGCCTCTGGAGGAA	1399
Db	181	AACCGCCTCTGGAGGTA	197

RESIST 1A

US-09-822-846-596
 ; Sequence 596, Application US/09822846
 ; Publication No. US20030027139A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Jacobs, Kenneth
 ; APPLICANT: McCoy, John M.
 ; APPLICANT: Lavallie, Edward R.
 ; APPLICANT: Collins-Racie, Lisa A.
 ; APPLICANT: Evans, Cheryl
 ; APPLICANT: Merberg, David
 ; APPLICANT: Treacy, Maurice
 ; APPLICANT: Agostino, Michael J.
 ; APPLICANT: Steininger II, Robert J.
 ; APPLICANT: Bowman, Michael R.
 ; APPLICANT: Spaulding, Vikki
 ; APPLICANT: Wong, Gordon G.
 ; APPLICANT: Clark, Hilary
 ; APPLICANT: Pechtel, Kim
 ; APPLICANT: Howes, Steven H.
 ; APPLICANT: Resnick, Richard J.
 ; APPLICANT: Gulkota, Kamalakar
 ; APPLICANT: Graham, James R.
 ; APPLICANT: Genetics Institute, Inc.
 ; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
 ; FILE REFERENCE: GIN 6400
 ; CURRENT APPLICATION NUMBER: US/09/822,846
 ; CURRENT FILING DATE: 2001-03-29
 ; PRIOR APPLICATION NUMBER: 60/195,605
 ; PRIOR FILING DATE: 2000-04-06
 ; NUMBER OF SEQ ID NOS: 629
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 596
 ; LENGTH: 3615
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-822-846-596

Db	241	TTTGTCTCCATCTCTTTCATTTCTCTGATGATCATTTTCCCTCGCATGGCTCGTCTTTTAT	300
QY	950	TTGTGTCGAAGGCTACGAATGCAAGAGCTCAAGACGAGGAAGCAGAGGCAATTAAGGCA	1009
Db	301	TACATCCAGAGGTTTCGATATGCAAAATGCCAGGATAGGAACCGAGCCGACCTGGGGGAT	360
QY	1010	GATGCTAAAAAGCTATTGGAAGGCTTCAACTACGCACACTGAAACAAGGAGACAAGGAA	1069
Db	361	GCAGCAAGAAAGCCATCAGCAACTCCAGATCAGGACCATCAAGAGGGTGACAAGGAA	420
QY	1070	ATTGGCCCTGATGGAGATAGTTGTGCTGTGTGCAATGAAATGTATATAACCAAAATGATTTG	1129
Db	421	ACAGAGTCTGATTTTGTGACAACTGTGCAGTTTGTATTTGAAGGGTACAAAGCCCAATGACGTT	480
QY	1130	GTAAGCATTTAACTGCGCAACCATATTTTCCATAGACATGTGTGACCCCATGGCTGTTA	1189
Db	481	GTCCGGATCCTGCCCTGCGGCATCTTTTCCACAAGTCTGTGTGTGACCCCTGGCTTCTA	540
QY	1190	AAACAACAAGCTTGGCCCATGTGCAAAATGTGACATCTCAAAGCTTTGGGAATTAGGTTG	1249
Db	541	GACCATCGTACCTGTCCCATGTGCAAGATGAACATTTCTTAAAGCCCTAGGATCCCGCC	600
QY	1250	GATGTTGA 1257	
Db	601	AATGCCGA 608	
RESULT 15			
US-09-935-390A-16			
; Sequence 16, Application US/09935390A			
; Patent No. US20020076761A1			
; GENERAL INFORMATION:			
; APPLICANT: Escobedo, Jaime			
; Oujianjin, Hu			
; Garcia, Pablo			
; Williams, Lewis T.			
; Kothakota, Srinivas			
; TITLE OF INVENTION: Secreted Human Proteins			
; NUMBER OF SEQUENCES: 38			
; CORRESPONDENCE ADDRESS:			
; ADDRESS: Chiron Corporation			
; STREET: 4560 Horton Street			
; CITY: Emeryville			
; STATE: CA			
; COUNTRY: USA			
; ZIP: 94608-2916			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Diskette			
; COMPUTER: IBM Compatible			
; OPERATING SYSTEM: DOS			
; SOFTWARE: FastSeq for Windows Version 2.0			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/09/935,390A			
; FILING DATE: 22-Aug-2001			
; CLASSIFICATION: <Unknown>			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: 08/988,671			
; FILING DATE: 1997-12-11			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Jane E. R. Potter			
; REGISTRATION NUMBER: 33,332			
; REFERENCE/DOCKET NUMBER: 1369,002			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (510) 923-2718			
; TELEFAX: (510) 655-3542			
; TELEX: <Unknown>			
; INFORMATION FOR SEQ ID NO: 16:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 1322 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: single			
; TOPOLOGY: linear			
; SEQUENCE DESCRIPTION: SEQ ID NO: 16:			

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 26, 2003, 00:58:15 ; Search time 57 Seconds
(without alignments)
9544.642 Million cell updates/sec

Title: US-09-854-300-7

Perfect score: 1774

Sequence: 1 agctggagctccacgcgggt.....ataggcaagtctctctcag 1774

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/prodata/1/ina/5A COMB.seq.*
2: /cgn2_6/prodata/1/ina/5B COMB.seq.*
3: /cgn2_6/prodata/1/ina/6A COMB.seq.*
4: /cgn2_6/prodata/1/ina/6B COMB.seq.*
5: /cgn2_6/prodata/1/ina/PCUS COMB.seq.*
6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	147.6	8.3	1253	2	US-08-786-606-6
2	77	4.3	912	2	US-09-090-567-3
3	76.4	4.3	2605	2	US-08-680-395-4
4	73.8	4.2	1378	1	US-08-075-533-20
5	73.8	4.2	1378	2	US-08-948-176-20
6	73.8	4.2	1378	5	PCT-US91-09160-20
7	73.8	4.2	1556	4	US-09-043-937A-3
8	73.8	4.2	1568	4	US-09-043-937A-1
9	73.2	4.1	1543	4	US-09-227-357-26
10	73	4.1	1545	4	US-09-227-357-125
11	73	4.1	1585	4	US-09-183-861-54
12	73	4.1	1585	4	US-09-022-765-54
13	73	4.1	2775	1	US-08-730-771-1
14	73	4.1	2775	4	US-09-060-208-1
15	73	4.1	3089	1	US-08-472-934-5
16	73	4.1	3089	2	US-08-323-460A-5
17	73	4.1	3089	2	US-08-461-146C-5
18	73	4.1	3089	3	US-08-461-145C-5
19	73	4.1	3089	4	US-08-628-829-9
20	73	4.1	3328	4	US-08-960-048-1
21	71.4	4.0	2045	4	US-08-795-088A-1
22	71	4.0	625	4	US-09-288-143-17
23	70.4	4.0	2085	2	US-08-668-128B-7
24	70.4	4.0	2085	2	US-08-905-445-7
25	69.4	3.9	633	4	US-08-818-112-6
26	69.4	3.9	633	4	US-08-818-111-6
27	69.4	3.9	633	4	US-09-056-556-6

Sequence 6, Appli
Sequence 101, App
Sequence 96, Appl
Sequence 101, App
Sequence 96, Appl
Sequence 1, Appli
Sequence 107, App
Sequence 4, Appli
Sequence 108, App
Sequence 14, Appl
Sequence 14, Appl
Sequence 5, Appli
Sequence 9, Appli
Sequence 6, Appli
Sequence 6, Appli
Sequence 1, Appli
Sequence 1, Appli

28 69.4 3.9 633 4 US-09-072-596-6
29 68.2 3.8 500 4 US-08-818-112-101
30 68.2 3.8 500 4 US-08-818-111-96
31 68.2 3.8 500 4 US-09-056-556-101
32 68.2 3.8 500 4 US-09-072-596-96
33 67.8 3.8 2418 4 US-09-285-379-1
34 66.6 3.8 2327 4 US-09-149-476-107
35 66.2 3.7 3023 4 US-09-308-022-4
36 66 3.7 752 4 US-08-976-259-108
37 66 3.7 849 3 US-08-646-538-14
38 66 3.7 849 4 US-09-503-222-14
39 66 3.7 1924 4 US-09-424-283-5
40 66 3.7 1958 4 US-09-215-221-9
41 66 3.7 2961 4 US-08-446-935-6
42 66 3.7 3699 3 US-08-646-538-6
43 66 3.7 3699 4 US-09-503-222-6
44 66 3.7 3792 2 US-08-992-334-1
45 66 3.7 3792 3 US-08-302-752-1

ALIGNMENTS

RESULT 1
US-08-786-606-6
; Sequence 6, Application US/08786606
; Patent No. 5861495
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Au-Young, Janice
; APPLICANT: Coleman, Roger
; APPLICANT: Goll, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN ZINC-BINDING
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/786,606
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy RJ
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0173 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1253 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-786-606-6

Query Match 8.3%; Score 147.6; DB 2; Length 1253;
Best Local Similarity 55.9%; Pred. No. 2.6e-31;
Matches 301; Conservative 0; Mismatches 234; Indels 3; Gaps 1;

```

QY 720 ATGAGTCATCCCATGCTCACCCTGGTGCAGTAGACATTTGTTGCAATCATGATCGCA 779
Db 242 AGGAGCCAGTTACCATGACTCATCCAGCACTGGAGATATTTATTTGCTGTATGATAACAG 301
QY 780 ATCTCAAAAGGCACAAAATTTCTGCAATCTATTCAAAGAGGCATACAAAGTGACATGTCA 839
Db 302 RATTGAGGGTAAGATATTTGAGTTATCTGGAGAAAACATCTCTGTACAAATGACAA 361
QY 840 TAGAAGTAGGGAAGAAAACATGGCCCTTGGG---TGAATCACTATTCAATTTTTTTTCTTT 896
Db 362 TAGCTGTGGAACCTCGAATGCCACCGAAGAACTTTCAGCCGTGGCTCTCTAGTCTTCGTGT 421
QY 897 CTGTCCTCTTTTATTTATTTAGGCGGCAACTGTGGGCTATTTTATCTTTTATTTCTGTC 956
Db 422 CAATATCTTTTATTTGTTTGTATTTATTTCTTCCAGCATGCTCATATTCTACTTCATTC 481
QY 957 GAAGGCTACGGAATCCAGAGCTCAAGCAGGAAGCAGAGCAATTAAGGCGAGATGCTA 1016
Db 482 AGAAGATCAGGTACAAAATGACGCGACAGAACCCGCTCGTCTCGAGATGCAGCA 541
QY 1017 ABAAGCTATTGGAAGGCTTCAACTACGCACCTGAAACAGGAGAGCAAGGAATTTGGCC 1076
Db 542 AGAAGCCATCAGTAAATTTGACAAACAGGACAGTAAAGAGGGTGACAGGAATCTGACC 601
QY 1077 CTGATGGAGATAGTTGTGTGTGCTGATTTGAATTTGATAAAACCAAAATGATTTGGTAGCCA 1136
Db 602 CAGACTTTGATCATTTGTGAGTCTGATAGAGAGCTATAAGCAGAAATGATGTGTCGAA 661
QY 1137 TCTTAAGTGAACCATATTTTCCATAGACATGTTGTGACCCATGGCTGTATAAACACA 1196
Db 662 TTCTCCCTGCAAGCATGTTTTCCAAATCCTGGTGAGATCCCTGGCTTAGTGAAACATT 721
QY 1197 AGACTTGGCCCATGTGCAATCTGCATATCTCAAAAGCTTTGGGAATTCAGGTGGATGT 1254
Db 722 GTACCTGTCTATGTGCAAACTTATATATTGAGGCGCTTGGGAATTTGCGGAATTT 779

RESULT 2
US-09-090-567-3
; Sequence 3, Application US/09090567
; Patent No. 5989549
; GENERAL INFORMATION:
; APPLICANT: Sullivan, Robert
; APPLICANT: Brub, Bruno
; APPLICANT: Lgar, Christine
; APPLICANT: Gaudreault, Christian
; TITLE OF INVENTION: Acrosomal Sperm Protein And
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swabey Ogilvy Renault
; STREET: 1600 - 1981 McGill College
; CITY: Montreal
; STATE: QC
; COUNTRY: Canada
; ZIP: H3A 2Y3
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/090,567
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy, Kevin P.
; REGISTRATION NUMBER: 26,674
; REFERENCE/DOCKET NUMBER: 13045-2"US" FC/CC

```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 514-845-7126
; TELEFAX: 514-288-8389
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 912 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; US-09-090-567-3

Query Match 4.3%; Score 77; DB 2; Length 912;
Best Local Similarity 94.1%; Pred. No. 1.le-11;
Matches 80; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AGCTGGAGCTCCACCGCGTGGCGCGCTCTAGAACTAGTGGATCCCCCGGGTGCAGG 60
Db 7 AGCTGGAGCTCCACCGCGTGGCGCGCTCTAGAACTAGTGGATCCCCCGGGTGCAGG 66
QY 61 AATTGGCACGAGCCGAGGAGCTGC 85
Db 67 AATTGGCACGAGCCGAGGAGCTGC 91

RESULT 3
US-08-680-395-4
; Sequence 4, Application US/08680395
; Patent No. 5892010
; GENERAL INFORMATION:
; APPLICANT: Gray, Joe W.
; APPLICANT: Collins, Colin
; APPLICANT: Hwang, Soo-in
; APPLICANT: Godfrey, Tony
; APPLICANT: Kowbel, David
; APPLICANT: Rommens, Johanna
; TITLE OF INVENTION: Genes from the 20q13 Amplicon and Their
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/680,395
; FILING DATE: 15-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-0689000US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2605 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY:

```

```
; LOCATION: 1..2605
; OTHER INFORMATION: /note= "cDNA clone cc43 of 4 kb
; OTHER INFORMATION: transcript"
US-08-680-395-4

Query Match 4.3%; Score 76.4; DB 2; Length 2605;
Best Local Similarity 80.9%; Pred. No. 2.8e-11;
Matches 89; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 AGCTGGAGCTCCACCGCGTGGCGCGCTCTAGAACTAGTGGATCCCCCGGCTGCAGG 60
Db 36 AGCTGGAGCTCCACCGCGTGGCGCGCTCTAGAACTAGTGGATCCCCCGGCTGCAGG 95

QY 61 AATTCCGACGACGCGGAGGAGTGCATCTCGGCAACTGTGTGCTGAAC 110
Db 96 AATTCCGACGACGCGGAGTGCATCTCGGCAACTGTGTGCTGAAC 145

RESULT 4
US-08-075-533-20
; Sequence 20, Application US/08075533
; Patent No. 5530186
; GENERAL INFORMATION:
; APPLICANT: Hitz, William D.
; APPLICANT: Yadav, Narendra S.
; TITLE OF INVENTION: Nucleotide Sequences of Soybean Acyl-ACP
; TITLE OF INVENTION: Thioesterase Genes
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. du Pont de Nemours and Company
; STREET: 1007 Market Street
; CITY: Wilmington
; STATE: Delaware
; COUNTRY: U.S.A
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/075,533
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/631,264
; FILING DATE: 20-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Morrissey, Bruce W.
; REGISTRATION NUMBER: 30,663
; REFERENCE/DOCKET NUMBER: CR-8926-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 992-4927
; TELEFAX: (302) 892-7949
; TELEX: 835420
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1378 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Brassica napus
US-08-075-533-20

Query Match 4.2%; Score 73.8; DB 1; Length 1378;
Best Local Similarity 97.4%; Pred. No. 1.1e-10;
Matches 75; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGCTGGAGCTCCACCGCGTGGCGCGCTCTAGAACTAGTGGATCCCCCGGCTGCAGG 60
Db 52 AGCTGGAGCTCCACCGCGTGGCGCGCTCTAGAACTAGTGGATCCCCCGGCTGCAGG 111
```

```
QY 61 AATTCCGACGACGCGGA 77
Db 112 AATTCCGACGACGAGA 128

RESULT 5
US-08-948-176-20
; Sequence 20, Application US/08948176
; Patent No. 5945585
; GENERAL INFORMATION:
; APPLICANT: Hitz, William D.
; APPLICANT: Yadav, Narendra S.
; TITLE OF INVENTION: ACYL-ACP THIOESTERASES GENES
; TITLE OF INVENTION: AND THEIR USE IN MATURING PLANT
; TITLE OF INVENTION: OIL COMPOSITION
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. du Pont de Nemours and Company
; STREET: 1007 Market Street
; CITY: Wilmington
; STATE: Delaware
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
; SOFTWARE: MICROSOFT WORD VERSION 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/948,176
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/631,264
; FILING DATE: DECEMBER 20, 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: CHRISTENBURY, LYNN M.
; REGISTRATION NUMBER: 30,971
; REFERENCE/DOCKET NUMBER: CR-8926-C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-5481
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1378 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Brassica napus
US-08-948-176-20

Query Match 4.2%; Score 73.8; DB 2; Length 1378;
Best Local Similarity 97.4%; Pred. No. 1.1e-10;
Matches 75; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGCTGGAGCTCCACCGCGTGGCGCGCTCTAGAACTAGTGGATCCCCCGGCTGCAGG 60
Db 52 AGCTGGAGCTCCACCGCGTGGCGCGCTCTAGAACTAGTGGATCCCCCGGCTGCAGG 111

QY 61 AATTCCGACGACGCGGA 77
Db 112 AATTCCGACGACGAGA 128

RESULT 6
PCT-US91-09160-20
; Sequence 20, Application PCT/US9109160
; GENERAL INFORMATION:
; APPLICANT: Hitz, William D.
; APPLICANT: Yadav, Narendra S.
```

;; TITLE OF INVENTION: Nucleotide Sequences of Soybean Acyl-ACP
;; TITLE OF INVENTION: Thioesterase Genes
;; NUMBER OF SEQUENCES: 22
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: E. I. du Pont de Nemours and Company
;; STREET: 1007 Market Street
;; CITY: Wilmington
;; STATE: Delaware
;; COUNTRY: U.S.A
;; ZIP: 19898
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US91/09160
;; FILING DATE: 19911216
;; CLASSIFICATION:
;; APPLICATION NUMBER: US 07/631,264
;; FILING DATE: 20-DEC-1990
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Morrissey, Bruce W.
;; REGISTRATION NUMBER: 30,663
;; REFERENCE/DOCKET NUMBER: CR-8926-A
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (302) 992-4927
;; TELEFAX: (302) 892-7949
;; TELEX: 835420
;; INFORMATION FOR SEQ ID NO: 20:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1378 base pairs
;; TYPE: NUCLEIC ACID
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; ORIGINAL SOURCE:
;; ORGANISM: Brassica napus
;; PCT-US91-09160-20

Query Match 4.2%; Score 73.8; DB 5; Length 1378;
Best Local Similarity 97.4%; Pred. No. 1.1e-10;
Matches 75; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGCTGGAGCTCCACCGCGTGGCGCGCTCTAGAACTAGTGGATCCCCCGGCTGCAGG 60
Db 52 AGCTGGAGCTCCACCGCGTGGCGCGCTCTAGAACTAGTGGATCCCCCGGCTGCAGG 111
QY 61 AATTCGGCAGAGCCGA 77
Db 112 AATTCGGCAGAGGA 128

RESULT 7
US-09-043-937A-3
; Sequence 3, Application US/09043937A
; Patent No. 6211432
; GENERAL INFORMATION:
; APPLICANT: BOUDET, ALAIN-MICHEL
; PICHON, MAGALIE
; GRIMA-PETTENATI, JACQUELINE
; BECKERT, MICHEL
; GAMAS, PASCAL
; BRIAT, JEAN-FRANCOIS
; TITLE OF INVENTION: DNA SEQUENCES CODING FOR CINNAMOYL-CoA
; REDUCTASE, AND APPLICATIONS THEREOF IN THE CONTROL OF
; LIGNIN CONTENTS IN PLANTS
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE, P.C.
; STREET: 110 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:

;; STATE: VA
;; COUNTRY: USA
;; ZIP: 22201
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/043,937A
;; FILING DATE: 24-Jul-1998
;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: WO PCT/FR96/01544
;; FILING DATE: 03-OCT-1996
;; APPLICATION NUMBER: FR 95.11623
;; FILING DATE: 03-OCT-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: SADOFF, B.J.
;; REGISTRATION NUMBER: 36,663
;; REFERENCE/DOCKET NUMBER: 1487-20
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 703-816-4000
;; TELEFAX: 703-816-4100
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1556 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA to mRNA
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 195..1310
;; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-043-937A-3

Query Match 4.2%; Score 73.8; DB 4; Length 1556;
Best Local Similarity 97.4%; Pred. No. 1.1e-10;
Matches 75; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGCTGGAGCTCCACCGCGTGGCGCGCTCTAGAACTAGTGGATCCCCCGGCTGCAGG 60
Db 52 AGCTGGAGCTCCACCGCGTGGCGCGCTCTAGAACTAGTGGATCCCCCGGCTGCAGG 111
QY 61 AATTCGGCAGAGCCGA 77
Db 112 AATTCGGCAGAGGA 128

RESULT 8
US-09-043-937A-1
; Sequence 1, Application US/09043937A
; Patent No. 6211432
; GENERAL INFORMATION:
; APPLICANT: BOUDET, ALAIN-MICHEL
; PICHON, MAGALIE
; GRIMA-PETTENATI, JACQUELINE
; BECKERT, MICHEL
; GAMAS, PASCAL
; BRIAT, JEAN-FRANCOIS
; TITLE OF INVENTION: DNA SEQUENCES CODING FOR CINNAMOYL-CoA
; REDUCTASE, AND APPLICATIONS THEREOF IN THE CONTROL OF
; LIGNIN CONTENTS IN PLANTS
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE, P.C.
; STREET: 110 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION NUMBER: US/09/043.937A
FILING DATE: 24-Jul-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR96/01544
FILING DATE: 03-OCT-1996
APPLICATION NUMBER: FR 95.11623
FILING DATE: 03-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 36.663
REFERENCE/DOCKET NUMBER: 1487-20
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1568 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 278..1306
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-043-937A-1

Query Match 4.2%; Score 73.8; DB 4; Length 1568;

Best Local Similarity 97.4%; Pred. No. 1.1e-10;
Matches 75; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGCTGAGTCCACCGCGTGGCGCGCTTAGAAGTAGTGATCCCGCGGTGCAGG 60
|||||
Db 47 AGCTGAGTCCACCGCGTGGCGCGCTTAGAAGTAGTGATCCCGCGGTGCAGG 106
|||||

QY 61 AATTCGGCAGGCGGA 77
|||||

Db 107 AATTCGGCAGGCGGA 123
|||||

RESULT 9

US-09-227-357-26
Sequence 26, Application US/09227357
Patent No. 6342581
GENERAL INFORMATION:
APPLICANT: Fischer et al.
TITLE OF INVENTION: 123 Human Secreted Proteins
FILE REFERENCE: P2010P1
CURRENT APPLICATION NUMBER: US/09/227,357
CURRENT FILING DATE: 1999-01-08
EARLIER APPLICATION NUMBER: PCT/US98/13684
EARLIER FILING DATE: 1998-07-07
EARLIER APPLICATION NUMBER: 60/051,926
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,793
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,925
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,929
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,803
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,732
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,931
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,932

EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,916
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,930
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,918
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,920
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,733
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,795
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,919
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,928
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/055,722
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,723
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,948
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,949
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,953
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,950
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,947
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,964
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/056,360
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,684
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,984
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,954
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/058,785
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,664
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,660
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,661
NUMBER OF SEQ ID NOS: 672
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 26
LENGTH: 1543
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (69)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (717)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (899)
OTHER INFORMATION: n equals a,t,g, or c
US-09-227-357-26

Query Match 4.1%; Score 73.2; DB 4; Length 1543;
Best Local Similarity 73.8%; Pred. No. 1.7e-10;
Matches 93; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 11 CCACCGGCTGGCGCCCTCTAGAACTAGTGGATCCCGGGCTGCAGGAATTCGGCAC 70
Db 99 CCACCGGCTGGCGCCCTCTAGAACTAGTGGATCCCGGGCTGCAGGAATTCGGCAC 158
QY 71 GAGCCGAGAGCTGATCTGGCGCAACCTGTGTCTGACGCTACGTGCTCTCTGGCTCCG 130
Db 159 GAGCCGACAGTAGGACATGTCATGGCATTCTTGTCTACCCCTTGTTCACCTCTCCGCCAG 218
QY 131 ACGTAG 136
Db 219 CCGTTG 224

RESULT 10
US-09-227-357-125
; Sequence 125, Application US/09227357
; Patent No. 6342581
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: P2010P1
; CURRENT APPLICATION NUMBER: US/09/227,357
; CURRENT FILING DATE: 1999-01-08
; EARLIER APPLICATION NUMBER: PCT/US98/13684
; EARLIER FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: 60/051,926
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,793
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,925
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,929
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,803
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,732
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,931
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,932
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,916
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,930
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,918
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,920
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,733
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,795
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,919
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,928
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/055,722
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,723
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,948
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,949
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,953
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,950
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,947
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,964
; EARLIER FILING DATE: 1997-08-18

; EARLIER APPLICATION NUMBER: 60/056,360
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,684
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,984
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,954
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/058,785
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,664
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,660
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,661
; EARLIER FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 125
; LENGTH: 545
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (7)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (16)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (41)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (42)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (87)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-227-357-125

Query Match 4.1%; Score 73; DB 4; Length 545;

Best Local Similarity 100.0%; Pred. No. 1e-10; Mismatches 0; Indels 0; Gaps 0;
Matches 73; Conservative 0;

QY 1 AGCTGGAGCTCCACCGGCTGGCGGCTCTAGAACTAGTGGATCCCGGGCTGCAGG 60
Db 117 AGCTGGAGCTCCACCGGCTGGCGGCTCTAGAACTAGTGGATCCCGGGCTGCAGG 176
QY 61 AATTCGGCAGG 73
Db 177 AATTCGGCAGG 189

RESULT 11

US-09-183-861-54
; Sequence 54, Application US/09183861
; Patent No. 6365165
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, David C.
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE THERAPY AND
; NUMBER OF SEQUENCES: 87
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle

STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/183,861
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/022,765
FILING DATE: 12-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.420C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 1585 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-183-861-54

Query Match 4.1%; Score 73; DB 4; Length 1585;
Best Local Similarity 100.0%; Pred. No. 1.9e-10;
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCTGGAGCTCCACCGGCTGTAGAACTAGTGGATCCCGGGCTGCAGG 60
|||||
DB 3 AGCTGGAGCTCCACCGGCTGTAGAACTAGTGGATCCCGGGCTGCAGG 62
QY 61 AATTCGGCAGG 73
|||||
DB 63 AATTCGGCAGG 75

RESULT 12
US-09-022-765-54
Sequence 54, Application US/09022765
Patent No. 6375955
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Campos-Neto, Antonio
APPLICANT: Webb, John R.
APPLICANT: Dillon, David C.
APPLICANT: Skeiky, Yasir A.W.
TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE THERAPY AND
NUMBER OF SEQUENCES: 87
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/022,765
FILING DATE: 12-FEB-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:

NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.420C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 1585 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-022-765-54

Query Match 4.1%; Score 73; DB 4; Length 1585;

Best Local Similarity 100.0%; Pred. No. 1.9e-10;
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCTGGAGCTCCACCGGCTGTAGAACTAGTGGATCCCGGGCTGCAGG 60
|||||
DB 3 AGCTGGAGCTCCACCGGCTGTAGAACTAGTGGATCCCGGGCTGCAGG 62
QY 61 AATTCGGCAGG 73
|||||
DB 63 AATTCGGCAGG 75

RESULT 13
US-08-730-771-1
Sequence 1, Application US/08730771
Patent No. 5780609
GENERAL INFORMATION:
APPLICANT: Marshall, Lisa
APPLICANT: Roehak, Amy
TITLE OF INVENTION: A NOVEL HUMAN B-CELL SURFACE MOLECULE
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-2799
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA: US/08/730,771
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Han, William T
REGISTRATION NUMBER: 34,344
REFERENCE/DOCKET NUMBER: P50380
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5219
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2775 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO

FRAGMENT TYPE:
ORIGINAL SOURCE:
US-08-730-771-1

Query Match 4.1%; Score 73; DB 1; Length 2775;
Best Local Similarity 100.0%; Pred. No. 2.6e-10; Indels 0; Gaps 0;
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGCTGGAGCTCCACCGCGGTGGCGCGCTCTAGAACTAGTGGATCCCCGGGCTGCAGG 60
Db 37 AGCTGGAGCTCCACCGCGGTGGCGCGCTCTAGAACTAGTGGATCCCCGGGCTGCAGG 96

Qy 61 AATTCGGCAGG 73
Db 97 AATTCGGCAGG 109

RESULT 14

US-09-060-208-1
Sequence 1, Application US/09060208
Patent No. 6221629

GENERAL INFORMATION:
APPLICANT: Marshall, Lisa
APPLICANT: Roshak, Amy
TITLE OF INVENTION: THE DNA SEQUENCE OF HUMAN RP-105
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:

ADDRESSEE: Ratner & Prestia
STREET: P.O. Box 980
CITY: Valley Forge
STATE: PA

COUNTRY: USA

ZIP: 19482-0980

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSEQ Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/060,208

FILING DATE: April 15, 1998

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Prestia, Paul F.

REGISTRATION NUMBER: 23,031

REFERENCE/DOCKET NUMBER: P50380-1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 610-407-0700

TELEFAX: 610-407-0701

TELEX: 846169

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2775 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE:

ORIGINAL SOURCE:

US-09-060-208-1

Query Match 4.1%; Score 73; DB 4; Length 2775;

Best Local Similarity 100.0%; Pred. No. 2.6e-10; Indels 0; Gaps 0;
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGCTGGAGCTCCACCGCGGTGGCGCGCTCTAGAACTAGTGGATCCCCGGGCTGCAGG 60
Db 37 AGCTGGAGCTCCACCGCGGTGGCGCGCTCTAGAACTAGTGGATCCCCGGGCTGCAGG 96

Qy 61 AATTCGGCAGG 73
Db 97 AATTCGGCAGG 109

RESULT 15

US-08-472-934-5

Sequence 5, Application US/08472934

Patent No. 5753446

GENERAL INFORMATION:

APPLICANT: JOHNSON, GARY L.

TITLE OF INVENTION: METHOD AND PRODUCT FOR REGULATING CELL

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lahive and Cockfield

STREET: 60 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/472,934

FILING DATE: 06-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/440,421

FILING DATE: 15-May-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/354,516

FILING DATE: 21-FEB-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 5,405,941

FILING DATE: 15-Apr-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/323,460

FILING DATE: 14-Oct-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/11690

FILING DATE: 14-Oct-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/04178

FILING DATE: 15-APR-1994

ATTORNEY/AGENT INFORMATION:

NAME: Giulio A. Deconti, Jr. Esq.

REGISTRATION NUMBER: 31,503

REFERENCE/DOCKET NUMBER: CPI-004DVCP2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400

TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 3089 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 400..2280

US-08-472-934-5

Query Match 4.1%; Score 73; DB 1; Length 3089;

Best Local Similarity 100.0%; Pred. No. 2.8e-10; Indels 0; Gaps 0;

Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGCTGGAGCTCCACCGCGGTGGCGCGCTCTAGAACTAGTGGATCCCCGGGCTGCAGG 60
Db 37 AGCTGGAGCTCCACCGCGGTGGCGCGCTCTAGAACTAGTGGATCCCCGGGCTGCAGG 96

Db 11 AGCTGGAGCTCCACCGGCTGGCGCGCTCTAGAACTAGTGGATCCCCCGGCTGCAGG 70
Qy 61 AATTGGCAGAG 73
Db 71 AATTGGCAGAG 83

Search completed: April 26, 2003, 02:28:24
Job time : 63 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 26, 2003, 00:56:15 ; Search time 1751 Seconds

(without alignments)

16408.220 Million cell updates/sec

Title: US-09-854-300-7

Perfect score: 1774

Sequence: 1 agctggagctccaccgggt.....ataggcaagtttctctcag 1774

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: gb_ges.*
18: em_ges_hum.*
19: em_ges_inv.*
20: em_ges_pln.*
21: em_ges_vrt.*
22: em_ges_fun.*
23: em_ges_mam.*
24: em_ges_mus.*
25: em_ges_other.*
26: em_ges_pro.*
27: em_ges_rtd.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1205.2	67.9	1921	11 AK018582	Mus muscu
2	1202	67.8	2741	11 AK004847	Mus muscu
3	836.8	47.2	1954	11 BC030951	Homo sapi
4	680.4	38.4	1806	11 BC012931	Homo sapi
5	673.8	38.0	811	13 B1600046	603251431
6	654.2	36.9	1000	13 B1600419	603246458

7	644.2	36.3	676	10 AV690200	AV690200
8	641	36.1	773	13 AK010196	AK010196
9	623.4	35.1	840	11 AK008312	Mus muscu
10	622.8	35.1	1150	13 B1488609	B1488609
11	622.4	35.1	906	12 BG185370	BG185370
12	622.2	35.1	683	10 AV686169	AV686169
13	615.4	34.7	996	12 BG176206	BG176206
14	577	32.5	681	13 B1220592	B1220592
15	564.4	31.8	683	10 AV692456	AV692456
16	542.8	30.6	557	13 BM507037	BM507037
17	537.2	30.3	904	12 BF664617	BF664617
18	531.2	29.9	596	12 BF231341	BF231341
19	524.2	29.5	862	12 BF668554	BF668554
20	515.6	29.1	902	12 BG165006	BG165006
21	513.8	29.0	841	12 BF696359	BF696359
22	509.8	28.7	812	13 B1110345	B1110345
23	502.2	28.3	872	12 BF699229	BF699229
24	499.4	28.2	789	12 BF696355	BF696355
25	496	28.0	879	12 BF668897	BF668897
26	492.4	27.8	759	12 BG400279	BG400279
27	482.4	27.2	774	12 BG432146	BG432146
28	481.6	27.1	841	12 BF666680	BF666680
29	481.6	27.1	847	12 BE958515	BE958515
30	475.6	26.8	804	13 B163069	B163069
31	471.8	26.6	856	12 BE958087	BE958087
32	469	26.4	680	10 BE308209	BE308209
33	468.6	26.4	884	12 BF542034	BF542034
34	466.4	26.3	835	12 BE958350	BE958350
35	461.4	26.0	622	12 BF137973	BF137973
36	459.2	25.9	725	13 B1557834	B1557834
37	458.6	25.9	466	12 BF382058	BF382058
38	456.4	25.7	732	12 BF246687	BF246687
39	443.6	25.0	813	13 B1648028	B1648028
40	442.6	24.9	619	13 BG991456	BG991456
41	441.8	24.9	889	12 BG172246	BG172246
42	440	24.8	862	12 BF700761	BF700761
43	435.4	24.5	854	12 BF670202	BF670202
44	431.4	24.3	759	12 BE958295	BE958295
45	427.2	24.1	942	12 BF967109	BF967109

ALIGNMENTS

RESULT 1	AK018582	1921 bp	mRNA	linear	HTC 19-JAN-2002
LOCUS	Mus musculus adult male cecum CDNA, RIKEN full-length enriched library, clone:9130001F19: BRAIN CDNA, CLONE MNCB-3816, SIMILAR TO AF171875 GI-RELATED ZINC FINGER PROTEIN (MUS MUSCULUS), full insert sequence.				
ACCESSION	AK018582				
VERSION	AK018582.1	GI:12858361			
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (strain: C57BL/6J) adult male cecum CDNA to mRNA, clone lib: RIKEN full-length enriched mouse cDNA library				
ORGANISM	Mus musculus				
REFERENCE	1				
AUTHORS	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99219253				
PUBMED	10349636				
REFERENCE	2				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				


```

QY 1392 TGAGGAAACAGTGCAGTCAACAAATGAAGTCTACAGCTGGTAAACCATGAAGCAAAATT 1451
Db 1317 TGAGGAAACAGTGCAGTCAACAAATGAAGTCTACAGCTGGTAAACCATGAAGCAAAATT 1376
QY 1452 CTGTGGCAGTGGATGTTATTCCTCATGTGACAAACCAACCTTTGAAGAGAGCAAACTC 1511
Db 1377 CTGTGGCAGTGGATGTTATTCCTCATGTGACAAACCAACCTTTGAAGAGAGCAAACTC 1436
QY 1512 CTAAATCAAGAGACTGCTGTTCGAGAAATTAATCTT-AAATCTGTGTAATAGAAACT 1570
Db 1437 CTAAATCAAGAGACTGCTGTTCGAGAAATTAATCTT-AAATCTGTGTAATAGAAACT 1496
QY 1571 TGAACCATTAG-TAATAACAGAACTGCCAATCAGGCGCTAGTTT-CTATTAATAAATGG 1628
Db 1497 TGAACCATTAG-TAATAACAGAACTGCCAATCAGGCGCTAGTTT-CTATTAATAAATGG 1556
QY 1629 ATAAATTAATAAATAAGAGTATCTAGAAAGTCTCAGATGACTAATAATTATGCTATA 1688
Db 1557 GTAAA---CGTAAACCAAGAAATGATCTGAAAGTCTGAGGTAACCTTATATTATATA 1613
QY 1689 GTTAAATGCTTTAAATAATTTAACTGTGTAACCTTTTTCACCAAACTCATTTATAATA 1748
Db 1614 GTTAAA--TGCTTAACATATTTACCCAGTACCGTTT--CCAAACCACTCACCATAACG 1669
QY 1749 TTTTTCATAGGCAAGTTTCCTCT 1771
Db 1670 TTTTTCATAGGCAAGTTTCCTCT 1692

RESULT 3
BC030951 1954 bp mRNA linear HTC 13-JUN-2002
LOCUS Homo sapiens, hypothetical protein FLJ23516, clone IMAGE.4280352,
DEFINITION mRNA.
ACCESSION BC030951
VERSION BC030951.1 GI:21410659
KEYWORDS HTC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1954)
Strausberg,R.
Direct Submission
Submitted (03-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-bhgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 41 Row: c Column: 10
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 13375696
This clone has the following problem: frame shifted.
FEATURES
Location/Qualifiers
1..1954
/organism="Homo sapiens"
/db_xref="LocusID:79589"
/db_xref="taxon:9606"

```

```

/clone="IMAGE:4280352"
/tissue type="Brain, primitive neuroectodermal"
/clone_lib="NIH MGC_56"
/lab_host="DH10B"
/notes="Vector: pDNR-LIB"
BASE COUNT 676 a 294 c 346 g 638 t
ORIGIN
Query Match 47.2%; Score 836.8; DB 11; Length 1954;
Best Local Similarity 98.7%; Pred. No. 2.3e-185;
Matches 864; Conservative 0; Mismatches 8; Indels 3; Gaps 2;
QY 900 TGTCTTTTTTATTTATTTACGCGCGCAACTGTGGGCTATTTATCTTTTATCTGCTCGAA 959
Db 1 TGTCTTTTTTATTTATTTACGCGCGCAACTGTGGGCTATTTATCTTTTATCTGCTCGAA 60
QY 960 GGCTACGGAATGCAAGAGCTCAAGCAGGAAGCAGAGGCAATTAAGGCGAGATGCTAAAA 1019
Db 61 GGCTACGGAATGCAAGAGCTCAAGCAGGAAGCAGAGGCAATTAAGGCGAGATGCTAAAA 120
QY 1020 AAGCTATTGGAAGGCTTCAACTACGCACACTGAAACCAAGAGACAAGGAAATTTGGCCCTG 1079
Db 121 AAGCTATTGGAAGGCTTCAACTACGCACACTGAAACCAAGAGACAAGGAAATTTGGCCCTG 180
QY 1080 ATGAGATAGTGTGTGTGTGCAATGAAATGTATATAACCAAAATGATTTGGTACGCACT 1139
Db 181 ATGAGATAGTGTGTGTGTGCAATGAAATGTATATAACCAAAATGATTTGGTACGCACT 240
QY 1140 TAACTGCAACCAATTTTCCATAAGACATGTGTTGACCCATGCTGTTAAACACACAGA 1199
Db 241 TAACTGCAACCAATTTTCCATAAGACATGTGTTGACCCATGCTGTTAAACACACAGA 300
QY 1200 CTTGCCCATGTGCAAAATGTGACATCTCAAAGCTTTGGGAATTTGAGGTGATTTGAAAG 1259
Db 301 CTTGCCCATGTGCAAAATGTGACATCTCAAAGCTTTGGGAATTTGAGGTGATTTGAAAG 360
QY 1260 ATGATCAGTGTCTTTACAGTCCCTGTATCCAATGAAATATCTAATAGTGCCTCTCC 1319
Db 361 ATGATCAGTGTCTTTACAGTCCCTGTATCCAATGAAATATCTAATAGTGCCTCTCC 420
QY 1320 ATGAAGAGTAATTCGAGGAGAGACCGCATCATCTGATATGCTTCAGTACAGGGAACAG 1379
Db 421 ATGAAGAGTAATTCGAGGAGAGACCGCATCATCTGATATGCTTCAGTACAGGGAACAG 480
QY 1380 ATGAACCGCTCTCTGGAGGAAACAGTGCAGTCAACAAATGAAGTCTACAGTGGTAAACC 1439
Db 481 ATGAACCGCTCTCTGGAGGAAACAGTGCAGTCAACAAATGAAGTCTACAGTGGTAAACC 540
QY 1440 ATGAAGCAAAATCTGTGGCAGTGGATGTTATCTCATGTGACAAACCAACCTTTGAG 1499
Db 541 ATGAAGCAAAATCTGTGGCAGTGGATGTTATCTCATGTGACAAACCAACCTTTGAG 600
QY 1500 AAGAGCAAACTCTTAATCAAGAGAGCTGCTGTTCGAGAAATTAATCTTAAATCTGTGTA 1559
Db 601 AAGAGCAAACTCTTAATCAAGAGAGCTGCTGTTCGAGAAATTAATCTTAAATCTGTGTA 660
QY 1560 AATGAAAACTTGAACCAATTAATTAACAGAACTGCCAATCAGGCGCTAGTTCTATTA 1619
Db 661 AATGAAAACTTGAACCAATTAATTAACAGAACTGCCAATCAGGCGCTAGTTCTATTA 720
QY 1620 ATAAATTCGATTAATTAATTAATTAAGTGATCTGAAAGTCTCAGATGACTTAATAT 1679
Db 721 ATAAATTCGATTAATTAATTAATTAAGTGATCTGAAAGTCTCAGATGACTTAATAT 780
QY 1680 TATGCTATAGTTAAATGGCTTAAAAATATTTAACTGTTAACTTTTTTCCACCAAACTC 1739
Db 781 TATGCTATAGTTAAA--TGGCTTAAAAATTTAACTGTTAACTTTTTTCCA-CAAACTC 837
QY 1740 ATTATATATTTTTCATAGGCAAGTTTCCTCTCAG 1774
Db 838 ATTATATATTTTTCATAGGCAAGTTTCCTCTCAG 872

```



```

Db 380 GTGAACCGTACGGTCTGGAGCTGAGCGAGGAGGCGGTGTACGCCAGGACTCGCCGCTG 439
Qy 494 GAGCCTGTGGCTGGGGTCTGTGTACCGCCGACCGGGCCGGGGCGCTTAAGCCCTGTAAAC 553
Db 440 GAGCCTGTGGCTGGGGTCTGTGTACCGCCGACCGGGCCGGGGCGCTTAAGCCCTGTAAAC 499
Qy 554 CCGCACAGCAATTTTCCACGGTGCACCGGTTTGGGGAAGCACCGGTGCAAGTCTCTTTGG-TT 612
Db 500 CCGCACAGCAATTTTCCACGGTGCACCGGTTTGGGGAAGCACCGGTGCAAGTCTCTTTGGTT 559
Qy 613 GGCCTCTATTCMAACCGCGCGGGGGGTGC-ACCTTCGCGAGACAAGATCATCTGCGCTTATG 671
Db 560 GGCCTCTATTCMAACCGCGCGGGGGGTGCACCTTCGCGAGACAAGATCATCTGCGCTTATG 619
Qy 672 AGAGATGGGCGTCTGGAGCGGTCACTTTTAACTTCCCGGAGACCGGCAATGAGGTCAATCC 731
Db 620 AGAGAGGGGGTCTGGAGCGGTCACTTTTAACTTCCCGGAGACCGGCAATGAGGTCAAT-C 678
Qy 732 CCATGTCACCGGGTGCAGTAGACATTGTTGCAATCATGATCGGCAATCTGAAGGCA 791
Db 679 CCATGTCACCGGGTGCAGTAGACATTGTTGCAATCATGATCGGCGCAATCTGAAGGG 738
Qy 792 CAAAATCTGCAATCTATTCAAGAGGCATACAGTACAAATGTCATAGAGTAGGGA 851
Db 739 CAAAATCTGCAATCTATTCAAGAG--CATACAGTGACCACCTGGTCTAGAGTACGGGA 796
Qy 852 AAAAAATCGCCCTTTG 868
Db 797 ACAACATGGCTGGGG 813

RESULT 7
AV690200
LOCUS AV690200 GK Homo sapiens cDNA clone GKCA02 5', mRNA sequence.
DEFINITION AV690200
ACCESSION AV690200
VERSION AV690200.1 GI:10292063
KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 676)
XU.X., Huang,J., Xu.Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
Hu,G., Gu,J., Chen,Z. and Han,Z.
Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
21625106
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers
1. .676
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GKCA02"
/tissue="liver"
/dev_stage="Adult"
/lab_host="SOLR"
/notes="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"
94 a 223 c 231 g 125 t 3 others

FEATURES
Source

```

```

Query Match 36.3%; Score 644.2; DB 10; Length 676;
Best Local Similarity 98.1%; Pred. No. 2.8e-140;
Matches 661; Conservative 0; Mismatches 11; Indels 2; Gaps 1;

```

```

Qy 74 CCGAGAGCTGCTATCGCGCAACCTGTGTCTGACGCTAGTGCTCTTCCACCTGGCGGCACTGCTCAAG 133
Db 4 CCGAGAGCTGCTATCGCGCAACCTGTGTCTGACGCTAGTGCTCTTCCACCTGGCGGCACTGCTCAAG 63
Qy 134 TAGCTCGAGCTCCCACTCTCACTCCATTCCTTCCACCTGGCGGCACTGCTCAAG 193
Db 64 TAGCTCGAGCTCCCACTCTCACTCCATTCCTTCCACCTGGCGGCACTGCTCAAG 123
Qy 194 ACCAGGGTCTCTGCCAAGCGCTAGGAGGGCGCTGTCAGGGGGCGCTAGGGAACCTGCGGAGC 253
Db 124 ACCAGGGTCTCTGCCAAGCGCTAGGAGGGCGCTGTCAGGGGGCGCTAGGGAACCTGCGGAGC 183
Qy 254 GCGCGGCCATCGGGCGCGCGCTCGGGCGCGGGTCTCTGCGCGGGTGGCTGGCGGCTTT 313
Db 184 GCGCGGCCATCGGGCGCGCGCTCGGGCGCGGGTCTCTGCGCGGGTGGCTGGCGGCTTT 243
Qy 314 TCCAGATTCTGCTGCGATGGTGTCTTCTGCTGGCCCTGAGTCCGCGAGGCACTGCGGCTTCCCGG 373
Db 244 TCCAGATTCTGCTGCGATGGTGTCTTCTGCTGGCCCTGAGTCCGCGAGGCACTGCGGCTTCCCGG 303
Qy 374 GGGGCTGAAGAGCTGTGGACCGCGTACCTCAACGTGTCTTGGCGGGTTCGCGACACGGGA 433
Db 304 GGGGCTGAAGAGCTGTGGACCGCGTACCTCAACGTGTCTTGGCGGGTTCGCGACACGGGA 363
Qy 434 GTGAACCGTACGGTGTGGAGCTGAGCGAGGAGGGCGGTGACGGCGAGGACTGCGCGCTG 493
Db 364 GTGAACCGTACGGTGTGGAGCTGAGCGAGGAGGGCGGTGACGGCGAGGACTGCGCGCTG 423
Qy 494 GAGCCTGTGGCTGGGGTCTCTGCTGACCGCCGACCGGGCGCGCTTAAACGCTGTAAAC 553
Db 424 GAGCCTGTGGCTGGGGTCTCTGCTGACCGCCGAGCGGGCCGCGGCGCTTAAACGCTGTAAAC 483
Qy 554 CCGCACAGCAATTTTACGGTGTGCGCGCGGGTTGGGGAAGCACCGGTGCAAGTCTCTTGGTTG 613
Db 484 CCGCACAGCAATTTTACGGTGTGCGCGCGGGTTGGGGAAGCACCGGTGCAAGTCTCTTGGTTG 543
Qy 614 GCCTCATCAACCGCGCGGGGCTGCACTTTCGAGACAAGATCCATCTGCGCTTATGAG 673
Db 544 GCCTCATCAACCGCGCGGGGCTGCACTTTCGAGACAAGATCCATCTGCGCTTATGAG 603
Qy 674 AGATGGCGCTGTGGAGCGCTCATCTTTAACTTCCCGGAGCCCGCAATGAGTCAATCCCC 733
Db 604 AGATGGCGCTGTGGAGCGCTCATCTTTTAACTTCCCGGAGCCCGCAATGAGTCAATCCCC 661
Qy 734 ATGCTCTACCCCGG 747
Db 662 ATGCTCTACCCCGG 675

```

```

RESULT 8
BI601096
LOCUS BI601096
DEFINITION 603249588F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:530380 5',
mRNA sequence.
ACCESSION BI601096
VERSION BI601096.1 GI:15494035
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 773)
XU.X., Huang,J., Xu.Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
Hu,G., Gu,J., Chen,Z. and Han,Z.
Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
21625106
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers
1. .676
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GKCA02"
/tissue="liver"
/dev_stage="Adult"
/lab_host="SOLR"
/notes="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"
94 a 223 c 231 g 125 t 3 others

BASE COUNT
ORIGIN

```

cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM11762 row: p column: 21
 High quality sequence stop: 759.

Location/Qualifiers
 1. 773

FEATURES

source
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5301380"
 /clone_lib="MHC MGC 96"
 /tissue_type="hypothalamus"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pBluescriptR (modified pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NHGRI/NIH), National Institutes of Health. Note: this is a NIH MGC library."
 BASE COUNT 130 a 242 c 254 g 146 t 1 others
 ORIGIN

Query Match 36.1%; Score 641; DB 13; Length 773;
 Best Local Similarity 96.3%; Pred. No. 1.6e-139;
 Matches 720; Conservative 0; Mismatches 21; Indels 7; Gaps 6;
 Qy 74 CCGAGGAGCTCATCTGGGCAACCTGTGTGCTGACGCTACGTCCTCTGGTCCGAGC 133
 Db |
 Qy 134 TAGCTCGAGTCCCGAGTCTCACTTCATTCCTTCCCACTCGGCGGCACTGCTCAAG 193
 Db |
 Qy 82 TAGCTCGAGTCCCGAGTCTCACTTCATTCCTTCCCACTCGGCGGCACTGCTCAAG 141
 Db |
 Qy 194 ACCAGGCTCTGCAAGCGCTAGGAGGCGGCTGCGGCGGCGCTAGGCACTGCGGAGC 253
 Db |
 Qy 254 GCGCGCGCATGGGCGCGCGCTGGGCGCGGGTCTCTGCGCGCGGTGGCTGGCGTTT 313
 Db |
 Qy 201 GCGCGCGCATGGGCGCGCGCTGGGCGCGGGTCTCTGCGCGCGGTGGCTGGCGTTT 260
 Db |
 Qy 314 TCCAGATTGCTGGCATGTGCTTCTGCTGCGCCTGAGTCGCGAGGACCGGTTCCCGG 373
 Db |
 Qy 261 TCCAGATTGCTGGCATGTGCTTCTGCTGCGCCTGAGTCGCGAGGACCGGTTCCCGG 320
 Db |
 Qy 374 GGGGCTGAAGCAGTGTGACCGCGTACCTCAACGCTGCTCTGCGGGTTCGCGACACGGGA 433
 Db |
 Qy 321 GGGGCTGAAGCAGTGTGACCGCGTACCTCAACGCTGCTCTGCGGGTTCGCGACACGGGA 380
 Db |
 Qy 434 GTGAACCGTAGCTGTGGAGCTGAGCGAGGCGGTGTACGCGCAGGACTCGCGCGTG 493
 Db |
 Qy 381 GTGAACCGTAGCTGTGGAGCTGAGCGAGGAGGCGGTGTACGCGCAGGACTCGCGCGTG 440
 Db |
 Qy 494 GAGCCTGTGGTGGGCTCTGTTACCGCGCGGCGCGGCGGCTTAACGCTGTAAAC 553
 Db |
 Qy 441 GAGCCTGTGGTGGGCTCTGTTACCGCGCGGCGCGGCGGCTTAACGCTGTAAAC 500
 Db |
 Qy 554 CGGCACACGAATTTTAC -GGTGGCCACCGGTTTGGGGAAGC -ACCGTGAAGTCTCTTGG- 610
 Db |
 Qy 501 CGGCACACGAATTTTACCGGTTTGGGGAAGC -ACCGTGAAGTCTCTTGGT 560
 Db |
 Qy 611 TTGGGCGCTCATCCACCGCGGCGGCGGTGACCTTCGCGAGACAGATCCATCTCGGCTAT 670
 Db |
 Qy 561 TTGGGCGCTCATCCACCGCGGCGGCGGTGACCTTCGCGAGACAGATCCATCTCGGCTAT 620
 Db |

Qy 671 GAGAGATGGCGTCTGGAGCGCTCATCTTTAACTTCCCGGAGCC--GCAATGAGGTCA 728
 Db |
 Qy 621 GAGAGATGGCGTCTGGAGCGCTCATCTTTAACTTCCCGGAGCCCGCAATGAGGTCA 680
 Db |
 Qy 729 TCCCATCTCTCACCCTGGGTCAGTAGACATT-GTTGCAATCATGCGCAATCTGAAA 787
 Db |
 Qy 681 TCCCATCTCTCACCCTGGGTCAGTAGACATTGTTGCAATCATGTCGCAATCTGCA 740
 Db |
 Qy 788 GGCACAAAATTTCTGCAATCTATTCAA 815
 Db |
 Qy 741 AGGCCAAAATTTCTGCAATCTATTCCAA 768
 Db |
 RESULT 9
 AK008312
 LOCUS
 DEFINITION
 Mus musculus adult male small intestine cDNA, RIKEN full-length enriched library, clone:2010104G21.BRAIN CDNA, CLONE MNCB-3816, SIMILAR TO AFI71875 GI-RELATED ZINC FINGER PROTEIN (MUS MUSCULUS), full insert sequence.
 ACCESSION
 AK008312
 VERSION
 AK008312.1 GI:12842420
 KEYWORDS
 HTC; CAP trapper.
 SOURCE
 Mus musculus (strain: C57BL/6J) adult male small intestine cDNA to mRNA, clone: libRIKEN full-length enriched mouse cDNA library clone:2010104G21.
 ORGANISM
 Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 REFERENCE
 1 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE
 99279253
 PUBMED
 10349636
 REFERENCE
 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 JOURNAL
 20499374
 MEDLINE
 11042159
 PUBMED
 REFERENCE
 3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 JOURNAL
 20530913
 MEDLINE
 11076861
 PUBMED
 REFERENCE
 4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuenli, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Badaralli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombarts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Sessa, T., Shibata, Y., Storch, K. F., Suzuki, H.,

source

```

1. 1150
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5191785"
/clone_lib="NIH_MGC_114"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb. Insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."
BASE COUNT 264 a 315 c 336 g 235 t
ORIGIN

Query Match 35.1%; Score 622.8; DB 13; Length 1150;
Best Local Similarity 95.8%; Pred. No. 3.1e-135;
Matches 747; Conservative 0; Mismatches 22; Indels 11; Gaps 10;

QY 174 CTGGCGGCACCTGCTCAAGACACAGGCTCTGCCAAGCGCTAGGAGCGCGCTGCCAGGG 233
DB 1 CTGGCGGCACCTGCTCAAGACCA-GTTCCTGCCAAGCGCTAGGAGCGCGCTGCCA-GG 58
QY 234 GCGCTAGGAACTCGGAGCGCGCGCGCATGGGGCCGCCCTGGGGCGGGGTCTCCT 293
DB 59 GCGCTAGGAACTCGGAGCGCGCGCGCATGGGGCCGCCCTGGGGCGGGGTCTCCT 118
QY 294 GCCGGGTGGTGGGCTTTTCCAGATTGCTGGCATGGTGTCTTCTGCTGGCCCTGATC 353
DB 119 GCCGGGTGGTGGGCTTTTCCAGATTGCTGGCATGGTGTCTTCTGCTGGCCCTGATC 178
QY 354 GCAGGCAACCGGTTTCCCG-GGGGGCTGAAGCAGTGTGGACCGCGTACCTCAACGTCTCC 412
DB 179 GCAGGCAACCGGTTTCCCGTGGGGCTGAAGCAGTGTGGACCGCGTACCTCAACGTCTCC 238
QY 413 TGGCGGGTTTCCGACACCGGAGTGAAACCGTACGTTGGAGTGTGAGCGAGGAGGGGTG 472
DB 239 TGGCGGGTTTCCGACACCGGAGTGAAACCGTACGTTGGAGTGTGAGCGAGGAGGGGTG 298
QY 473 TACGCCAGGACTCGCCGCTGGAGCCTGTGCTGGGGTCTGGTACCGCCGCGGAGGCC 532
DB 299 TACGCCAGGACTCGCCGCTGGAGCCTGTGCTGGGGTCTGGTACCGCCGCGGAGGCC 358
QY 533 GGGCGGCTTAAC-GCCTGTAAACCGGACACAGAAATTTCACGTTGCCCGCTTGGGGAAG 591
DB 359 GGGCGGCTTAACGCTGTAAACCGGACACAGAAATTTCACGTTGCCCGCTTGGGGAAG 418
QY 592 CACCGTGAAGTCTCTTGGTTGGCCCTCATCAACGCGCGCGGGGCTGCACCTTCGAG 651
DB 419 CACCGTGAAGTCTCTTGGTTGGCCCTCATCAACGCGCGCGGGGCTGCACCTTCGAG 478
QY 652 C-AAGATCATCTGCTTTATGAGAGTGGGGCTCTGGAGCGGTCACTTTAACTTCCCG 710
DB 479 CGAAGTCCATCTGCTTTATGAGAGGAGGGGCTCTGGAGCGGTCACTTTAACTTCCCG 538
QY 711 GGACCGCAATGAGGTCACTCCCATGTCTCAACCGGGTGCAGTGAACATTTGCAATCA 770
DB 539 GGACCGCAATGAGGTCACTCCCATGTCTCAACCGGGTGCAGTGAACATTTGCAATCA 598
QY 771 TGATCGCAATCTGAAGGCAAC-AAAAATTCGCAATCTATTCAAGAGGCATCAAGTG 829
DB 599 TGATCGCAATCTGAAGGCAACAAAAATTCGCAATCTATTCAAGAGGCATCAAGTG 658
QY 830 ACAATGGTCATAGAAGTA-GGGAAAAAATCATGGCCCTTTGGGTGAATC-ACTATTCAATTT 887
DB 659 ACAATGGTCATAGAAGTACGGGAAAAAATCATGGCCCTTCGGGTGAATCACTATTCAATA 718
QY 888 TTTTGGTTTC--TGTTGCTTTTATTTA-TTACGGCGGCAACTGTGGGCTATTATTCT 944
DB 719 TTTTGGTAACCTGGTGTCTTCCCATTTACTTTACGGCGGCAACTGTGGGCTATTATTGTAT 778

```

RESULT 11

```

BG165370 602346005P1 NIH_MGC_89 Homo sapiens cDNA clone IMAGE:4456033 5',
LOCUS mRNA sequence.
DEFINITION BG165370.1 GI:12672073
ACCESSION BG165370
VERSION BG165370.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 906)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10250 row: j column: 02
High quality sequence stop: 649.

FEATURES
source
1. 906
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4456033"
/clone_lib="NIH_MGC_89"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: kidney; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 275 a 204 c 183 g 244 t
ORIGIN

Query Match 35.1%; Score 622.4; DB 12; Length 906;
Best Local Similarity 96.0%; Pred. No. 3.7e-135;
Matches 692; Conservative 0; Mismatches 22; Indels 7; Gaps 5;

QY 1058 GGAGACAAGGAATTTGGCCCTGATGGAGATAGTTGCTGTGTCGATTTGAATTGTATAA 1117
DB 1 GGAGACAAGGAATTTGGCCCTGATGGAGATAGTTGCTGTGTCGATTTGAATTGTATAA 59
QY 1118 CAAATGATTTGGTAGCATCTTAACGTGCAACCATATTTTCCATAAGACATGTGTGAC 1177
DB 60 CCAATGA-TTGGTAGCATCTTAACGTGCAACCATATTTTCCATAAGACATGTGTGAC 118
QY 1178 CATGTGCTTTAAACACAAAGACTTGGCCCATGTGCAATGTGACATCTCAAGCTTTG 1237
DB 119 CATGTGCTTTAAACACAAAGACTTGGCCCATGTGCAATGTGACATCTCAAGC-TTG 177
QY 1238 GGAATTTGAGTGGATGTTTGAAGATGATCAGTGTCTTTTCAAGTCCCTGTATCCAATGA 1297
DB 178 GGAATTTGAGTGGATGTTTGAAGATGATCAGTGTCTTTTCAAGTCCCTGTATCCAATGA 237
QY 1298 ATATCTAATAGTGCCTCTCCCATGAAGAGGATAATTCGAGCGAGACCGCATCATCTGA 1357
DB 238 ATATCTAATAGTGCCTCTCCCATGAAGAGGATAATTCGAGCGAGACCGCATCATCTGA 297
QY 1358 TATGCTTCAGTACAGGGAACAGATGATGAAACCGCTCTGGAGGAACACGTCAGTCAAAAT 1417
DB 298 TATGCTTCAGTACAGGGAACAGATGATGAAACCGCTCTGGAGGAACACGTCAGTCAAAAT 357

```

	Best Local Similarity	97.0%; Pred. No. 4e-135;			
	Matches 655; Conservative	0; Mismatches 18; Indels	2; Gaps		
Qy	74	CCGAGGAGCTGCATCTGCGCAACCTGTGTGCTGACGTACGTGCTCTCTGGCTCCGACG	133		
Db	4	CCGAGGAGCTGCATCTGCGCAACCTGTGTGCTGACGTACGTGCTCTCTGGCTCCGACG	63		
Qy	134	TAGTCGCAGCTCCCCAGTCTCACTCCATTCCTTCCCACTCTGGCGCGCACTGCTCAAG	193		
Db	64	TAGTCGCAGCTCCCCAGTCTCACTCCATTCCTTCCCACTCTGGCGCGCACTGCTCAAG	123		
Qy	194	ACAGGGTCTGCCAAAGCGTTAGAGGGCGCGTGCACAGGGCGCTAGGNACTGCGGAGC	253		
Db	124	ACAGGGTCTGCCAAAGCGTTAGAGGGCGCGTGCACAGGGCGCTAGGNACTGCGGAGC	183		
Qy	254	GCAGCGCCATGGGGCGCGCGCTGGGGCGGGGTCTCTTCGCGCGTGGCTCGCGGCTTT	313		
Db	184	GCAGCGCCATGGGGCGCGCGCTGGGGCGGGGTCTCTTCGCGCGTGGCTCGCGGCTTT	243		
Qy	314	TCCAGATTGCTGGCATGGTGCTTCTCTGCTGGCCCTGAGTCCGAGGACCCCGTTCCCGG	373		
Db	244	TCCAGATTGCTGGCATGGTGCTTCTCTGCTGGCCCTGAGTCCGAGGACCCCGTTCCCGG	303		
Qy	374	GGGCTGAAGCAGTGTGGACGGGGTACCTCAACGTGTCTGGCGGGTTCCGCACACGGGA	433		
Db	304	GGGCTGAAGCAGTGTGGACGGGGTACCTCAACGTGTCTGGCGGGTTCCGCACACGGGA	363		

364	GTGAACCGTACCGGTGTGGAGCTGAGCGAGAGGGGTGTACGGCANGACTCCGCGTG	553b
Db		
494	GAGCCTCTGGCTGGGTCTCTGGTACCGCCGACGGGCGCGGGGGCGTTTACCGCTGTAAAC	553c
Qy		
424	GAGCCTCTGGCTGGGTCTCTGGTACCGCCGACGGGCGCGGGTGTACCGCTGTAAAC	483b
Db		
554	CCGCACACGAAATTTCACGGTGGCCACGGTTTGGGAGACACCGTGCAGGTCTCTTGTTG	613b
Qy		

Db	484	CCGCACACGAATTTCA	CGGTGTCCACACGGT	TTGGGGAAAGCACCG	TGCAGATCTCTTTGTG	543
Qy	614	GCCCTCATCCAAACG	CGCGGGGGGTGCA	CTTTCGCACAAGAT	TCCATCTGGGTTATGAG	673
Db	544	GCCCTCATTTCAA	CGCCCGGGGGGTG	CACTTTCGCAGAC	CCAGATACATCTTGCTTATGAG	603
Qy	674	-AGATGGCGCTCTG	GAGCCGCTCATCTT	TAACTTCCCGGGAC	CCGCAATGAGGTCATCCC	733
Db	604	AAGAAGGGCGTCTG	AGCCCGCATCTTT	ACTTTCCCGGG	-CCCCCAATGAGGTCATCC	666
Qy	733	CATGCTCTACCCGGG	747			
Db	663	CATGTTTATCCCGGG	677			

RESULT 13	BG176206	BG176206	996 bp	linear	EST 06-FEB-
LOCUS	BG176206	BG176206	996 bp	linear	
DEFINITION	602335351P2 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:44582525				
ACCESSION	mRNA sequence.				
	BG176206	BG176206			

VERSION BGI76206.1 GI:12682909
 KEYWORDS
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; 1 (cases 1 to 1996)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.

CONTACT: ROBERT STRAUSSBERG, F.R.D.
Email: cgabs-r@math.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrived by: The I.M.A.G.E. Consortium (LLNL)

Query Match 35.1%; Score 622.2; DB 10; Length 683;

DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM10256 row: f column: 13
 High quality sequence stop: 668.
 Location/Qualifiers

FEATURES

source
 1. .996
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4458252"
 /clone_lib="NCI_CGAP_Mam1"
 /cissue_type="tumor, biopsy sample"
 /dev_stage="P0 months, virgin"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: Sali;
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"
 294 a 214 c 243 g 244 t 1 others
 BASE COUNT 294 a 214 c 243 g 244 t 1 others
 ORIGIN
 Query Match 34.7%; Score 615.4; DB 12; Length 996;
 Best Local Similarity 87.7%; Pred. No. 1.7e-133;
 Matches 717; Conservative 0; Mismatches 97; Indels 4; Gaps 4;
 QY 773 ATGGCAATCTGAAGGCAAAATTTCTGCAATCTATTCAAGAGGCATCAAGTGACA 832
 Db 1 ATTGGCAATCTGAAGGCAAAATTTCTGCAATCTATTCAAGAGGCATCAAGTGACA 60
 QY 833 ATGGTCATAGAGTGGGAAACATGGCCCTGGTGGTGAATCACTATTCAATTTTTC 892
 Db 61 ATGGTCATAGAGTGGGAAACATGGCCCTGGTGGTGAATCACTATTCAATTTTTC 119
 QY 893 GTTCTGTGTCCTTTTATTATTAGCGGCAACTGTGGCTATTATCTTTTATTCT 952
 Db 120 GTTCTGTGTCCTTTTATTATTAGCGGCAACTGTGGCTATTATCTTTTATTCT 179
 QY 953 GCTCGAAGGCTACGAAATGCAAGAGCTCAAGCAGGAAGCAGGCAATTAAGGCAGAT 1012
 Db 180 GCTCGAAGATTAAGAAATGCAAGAGCTCAAGCAGGAAGCAGGCAATTAAGGCAGAT 239
 QY 1013 GCTAAAGAAAGCTATTGGAAGCTTCAACTAGCAGCACTGAAACAGGAGCAGGAAAT 1072
 Db 240 GCTAAAGAAAGCTATTGGAAGCTTCAAGTGGCACTTGAACAGGAGCAGGAAAT 299
 QY 1073 GGCCCTGATGGAGATAGTTGTGTGTGCAATTTGTAATTTGTAACCAATGATTTGGTA 1132
 Db 300 GGCCCTGATGGAGATAGTTGTGTGTGCAATTTGTAATTTGTAACCAATGATTTGGT 359
 QY 1133 CGCATCTTAAGCTGACACCATATTTTCCATAGACATGTTGACCCATGGCTTTAAAA 1192
 Db 360 CGCATCTTAAGCTGACACCATATTTTCCATAGACATGTTGACCCATGGCTTTAGAA 419
 QY 1193 CACAAGACTTGGCCCATGTGCAATGTGACATCTCAAGCTTTGGGAATTTGAGTGGAT 1252
 Db 420 CACAAGACTTGGCCCATGTGCAATGTGACATCTCAAGCTTTGGGAATTTGAGTGGAT 479
 QY 1253 GTTGAAGATGATGATGTTCTTTACAGTCCCTGTATCCAAATGAAATCTAATAGTGC 1312
 Db 480 GTTGAAGATGATGATGTTCTTTACAGTCCCTGTATCCAAATGAAATCTAATAGTGC 539
 QY 1313 TCCTCCCATGAAGAGGATATCCAGGAGCCGACATCATCTGGATATGCTTCAGTACAG 1372
 Db 540 TCCTCCCATGAAGAGGACATCCAGTGGACATGATCTGGATATGCTTCAGTACAG 599
 QY 1373 GGAAACAGATGAACCGCTCTGGAGGAAACAGTGCAGTCAACAAATGAAAGTCTTACAGCTG 1432
 Db 600 GGAAACAGATGAACCGCTCTGGAGGAAACAGTGCAGTCAACAAATGAAAGTCTTACAGCTG 659
 QY 1433 GTAAACCATGAAGCAAAATCTGTGGCAGTGGATGTTATTCCTCATGTNGAC-AAACCAAC 1491

Db 660 GT-AACCATGAAGCAAAATCTGTGGCCGTGGATGTGTCCCATGTGTGACAAACCAAC 718
 QY 1492 CTTTGAAG-AAGACGAATCTTAATCAAGAGACTGTGTTCGAGAAATTAATCTTAAA 1550
 Db 719 CTTTGAAGAAAGATGAATCCCTGATCAAGAGGAGTTGTCCGGAGATTCAGTCTTACA 778
 QY 1551 ATCTGTGTAAATGAAACCTTGAAACCATTAAGTAATAAC 1588
 Db 779 ATCTGTGTCCATTACAAAATTAACCGTGTGTACACAC 816

RESULT 14
 B1220592
 LOCUS 602936079F1 NCI_CGAP_Li9 Mus musculus cDNA clone IMAGE:5099346 5',
 DEFINITION mRNA sequence.
 B1220592
 VERSION B1220592.1 GI:14674036
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 681)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM11238 row: n column: 19
 High quality sequence stop: 681.
 Location/Qualifiers

FEATURES

source
 1. .681
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:5099346"
 /clone_lib="NCI_CGAP_Li9"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: Sali; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.9 kb. Constructed by Life
 Technologies. Note: this is a NCI_CGAP Library."
 BASE COUNT 167 a 165 c 193 g 156 t
 ORIGIN

Query Match 32.5%; Score 577; DB 13; Length 681;
 Best Local Similarity 92.2%; Pred. No. 1.6e-124;
 Matches 630; Conservative 0; Mismatches 50; Indels 3; Gaps 2;
 QY 441 GTACGGTGTGGAGCTGAGCGAGGAGGCGGTGTACGCCAGGACTCCCGCTGGAGCTG 500
 Db 1 GCACGGTGTGGAGCTGAGCGAGGAGGCGGTGTACGCCAGGACTCCCGCTGGAGCTG 60
 QY 501 TGGCTGGGGTCTCTGGTACCGCCCGACCGGGCCGGCGCTTAACGCTGTAAACCCGACA 560
 Db 61 TCTCGGGGTCTCTGGTACCGCCCGACCGGGCCGGCGCTTAACGCTGTAAACCCGACA 120
 QY 561 CGAATTTACGGTGTCCCAACCGTTTGGGAAAGCACCGGTGCAAGTCTCTTGGTGGCCCTCA 620
 Db 121 CCAATTTACGGTGTCCCAACCGTTTGGGAAAGCACCGGTGCAAGTCTCTTGGTGGCCCTCA 180
 QY 621 TCCAAACCGGGGGGTGCACTTCCGACACAGATCCATCTGCTTATGAGAGATGGG 680
 Db 181 TCCAAACCGGGGGGTGCACTTCCGACACAGATCCATCTGCTTATGAGAGATGGG 240


```

Qy 681 CGTCTGAGCGTCATCTTTAACTTCCCGGACCGCGCAATGAGGTGATCCCATGCTCTC 740
Db 241 CTTCTGGAGCGGTGATCTTTAACTTCCCGGACCGCGCAATGAGGTGATCCCATGCTCTC 300
Qy 741 ACCCGGTGAGTAGACATGTTGCAATCATGATCGGCAATCTGAAAGGCACAAAATTC 800
Db 301 ACCCGGTGAGTAGACATGTTGCAATCATGATCGGCAATCTGAAAGGCACAAAATTC 360
Qy 801 TGAATCTATTCAAGAGGCATACAGTGCATGCAATGCTAGAGTAGGGAACAAACATG 860
Db 361 TGCAGTCTATTCAAGAGGCATACAGTGCATGCAATGCTAGAGTAGGGAACAAACATG 418
Qy 861 GCCCTTGGGTG-AATCACTATTCAATTTTTTCTGTTCTGTCCTCTTTTATTATTAGG 919
Db 419 GCCCTTGGGTGAAATCATATTCAATTTTTTCTGTTCTGTCCTCTTTTCAATATTAG 478
Qy 920 GCGCAACTGTGGGTATTTTATCTTTTATCTGTCGAAAGGCTAGCGAATGCAAGACT 979
Db 479 GCAGCAACCGTGGGTATTTTATCTTTTATCTGTCGAAAGGCTAGCGAATGCAAGACT 538
Qy 980 CAAAGCAGGAGCAGAGGCAATTAAGGCGAGATGCTAAAGAAAGCTATTGGAAGGCTTCAA 1039
Db 539 CAAAGCAGGAGCAGAGGCAATTAAGGCGAGATGCTAAAGAAAGCTATTGGAAGGCTTCAA 598
Qy 1040 CTACGCACACTGAAACAAGGAGACAAGAAATTTGGCCCTGATGAGATAGTTGTGCTGTG 1099
Db 599 CTGCGCACCTTGAACAAGGAGACAAGAAATTTGGCCCTGATGAGATAGTTGTGCTGTG 658
Qy 1100 TGCATTGAATGTATAACCAAA 1122
Db 659 TGCATTGAGCTCTATAAGCCAAA 681

RESULT 15
LOCUS AV692456 683 bp mRNA linear EST 16-JAN-2002
DEFINITION AV692456 GK Homo sapiens cDNA clone KCCBD09 5', mRNA sequence.
ACCESSION AV692456
VERSION AV692456.1 GI:10294319
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Xiao, H., Qu, J., Liu, F., Huang, Q., Cheng, Z., Li, N., Du, J., Hu, W.,
Shen, K., Lu, G., Fu, G., Zhong, M., Xu, S., Gu, W., Huang, W., Zhao, X.,
Hu, G., Gu, J., Chen, Z. and Han, Z.
Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
21625106
Contact: Zequang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers
1. .683
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GKCCBD09"
/clone_lib="GKC"
/tissue_type="hepatocellular carcinoma"
/dev_stage="Adult"
/lab_host="SOLR"
/note="vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

```

```

BASE COUNT 225 a 127 c 126 g 200 t 5 others
ORIGIN
Query Match 31.8%; Score 564.4; DB 10; Length 683;
Best Local Similarity 98.0%; Pred. No. 1.4e-121;
Matches 591; Conservative 0; Mismatches 9; Indels 3; Gaps 2;
Qy 1172 GTTCACCATGGCTGTTTAAACACACAGACTTTGCCCATGTGCAAAATGTGCATACTCAAA 1231
Db 1 GTTGACCATGGCTGTTTAAACACACAGACTTTGCCCATGTGCAAAATGTGCATACTCAAA 60
Qy 1232 GCTTTGGGAATGAGGTGGATGTTGAAGATGAGTCAAGTGTCTTTTACAAGTCCGTATCC 1291
Db 61 GCTTTGGGAATGAGGTGGATGTTGAAGATGAGTCAAGTGTCTTTTACAAGTCCGTATCC 120
Qy 1292 AATGAATATCTAATAGTCTCTCCCATGAAGAGATTAATCGACGAGACCGCATCA 1351
Db 121 AATGAATATCTAATAGTCTCTCCCATGAAGAGATTAATCGACGAGACCGCATCA 180
Qy 1352 TCTGGATATGCTTCAGTACAGGGAACAGATGAACCGCTCTGGAGGAACACGTCAGTCA 1411
Db 181 TCTGGATATGCTTCAGTACAGGGAACAGATGAACCGCTCTGGAGGAACACGTCAGTCA 240
Qy 1412 ACAATGAAGTCTACAGCTGGTAAACCATGAAGCAAAATCTGTGGCAGTGGATGTTATT 1471
Db 241 ACAATGAAGTCTACAGCTGGTAAACCATGAAGCAAAATCTGTGGCAGTGGATGTTATT 300
Qy 1472 CCTCATGTNGACAACCCAACTTTGAAGAGACGAAACTCTCTAATCAAGAGACTGCTGTT 1531
Db 301 CCTCATGTNGACAACCCAACTTTGAAGAGACGAAACTCTCTAATCAAGAGACTGCTGTT 360
Qy 1532 CGAGAAATTAATCTTAAATCTGTGTAATAAGAAAACTTTGAACCATTTAGTAATAACAGA 1591
Db 361 CGAGAAATTAATCTTAAATCTGTGTAATAAGAAAACTTTGAACCATTTAGTAATAACAGA 420
Qy 1592 ACTGCCAATCAGGCGCTAGTTTCTTAAATAAATGGATAAATTTAATAAATGAAGTG 1651
Db 421 ACTGCCAATCAGGCGCTAGTTTCTTAAATAAATGGATAAATTTAATAAATGAAGTG 480
Qy 1652 ATACTGAAAGTGTCTCAGATGACTAATATTATGCTATAGTTAAAAATGGCTTAAAAATATT 1711
Db 481 ATACTGAAAGTGTCTCAGATGACTAATATTATGCTATAGTTAAAAATGGCTTAAAAATATT 538
Qy 1712 AACCTGTTAACTTTTTTCCACCAAACTCATTATTAATTTTTTATAGGCAAGTTTCTCT 1771
Db 539 AACCTGTTAACTTTTTTCCCA-CAAACTCATTATAATATTTTTTCATAGGCAAGTTTCTCT 597
Qy 1772 CAG 1774
Db 598 CAG 600

```

Search completed: April 26, 2003, 02:27:20

Job time : 1767 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 08:27:01 ; Search time 42 Seconds
(without alignments)
816.561 Million cell updates/sec

Title: US-09-854-300-8

Perfect score: 2275

Sequence: 1 MGPPGAGVSCRGCGFSRL.....PTFEDETPNQETAVERIKS 428

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1134	49.8	225	10	US-09-764-864-850
2	838	36.8	165	10	US-09-764-864-1306
3	656.5	28.9	400	10	US-09-935-390A-20
4	468.5	20.6	276	10	US-09-935-390A-35
5	396	17.4	147	10	US-09-764-864-1272
6	246.5	10.8	381	9	US-09-798-889-65
7	227.5	10.0	296	10	US-09-925-300-1231
8	223	9.8	259	9	US-09-798-889-127
9	137.5	6.0	337	10	US-09-828-303-18
10	130.5	5.7	199	10	US-09-764-864-1291
11	130.5	5.7	207	10	US-09-925-302-638
12	130	5.7	461	10	US-09-764-864-1288
13	130	5.7	636	10	US-09-764-864-832
14	129.5	5.7	604	9	US-09-764-864-836
15	127.5	5.6	664	9	US-10-176-847-40
16	127.5	5.6	664	10	US-09-898-533-2
17	126.5	5.6	685	10	US-09-745-763-11
18	123.5	5.4	291	10	US-09-764-864-1293
19	118.5	5.2	171	10	US-09-764-864-840

20	118.5	5.2	171	10	US-09-764-864-1295
21	105.5	4.6	124	10	US-09-764-864-803
22	105.5	4.6	124	10	US-09-764-864-1264
23	105	4.6	155	10	US-09-925-300-1159
24	105	4.6	205	10	US-09-764-864-1260
25	105	4.6	224	10	US-09-764-864-799
26	103.5	4.5	662	10	US-09-799-777-53
27	103.5	4.5	1863	9	US-09-734-672-2
28	103.5	4.5	1863	9	US-09-734-672-4
29	103.5	4.5	1863	9	US-09-734-672-6
30	103.5	4.5	1863	9	US-09-982-828-2
31	103.5	4.5	1863	9	US-09-982-828-4
32	103.5	4.5	1863	9	US-09-982-828-6
33	103.5	4.5	1863	9	US-10-022-819-2
34	102.5	4.5	426	10	US-09-764-864-809
35	102.5	4.5	664	10	US-09-780-525-2
36	102	4.5	26926	9	US-09-759-508B-2
37	101.5	4.5	825	9	US-10-121-988-161
38	101.5	4.5	826	9	US-10-121-988-47
39	101.5	4.5	826	10	US-09-894-998-47
40	100	4.4	600	10	US-09-764-864-1282
41	100	4.4	615	10	US-09-925-301-1094
42	100	4.4	759	10	US-09-764-864-823
43	98.5	4.3	439	10	US-09-764-864-849
44	97	4.3	266	10	US-09-764-864-834
45	96.5	4.2	811	10	US-09-840-787-45

ALIGNMENTS

RESULT 1
US-09-764-864-850
; Sequence 850, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 850
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-864-850

Query Match	49.8%	Score 1134;	DB 10;	Length 225;
Best Local Similarity	99.1%	Pred. No. 5.9e-106;		
Mismatches	218;	Conservative	2;	Mismatches 0;
Indels	0;	Gaps	0;	

QY	209	FFVSVSFFIITAATGVYFIFYSARRLRNARQSRKQRLKADAKKAIQRLQLRTLKQDK	268
Db	6	FFVSVSFFIITAATGVYFIFYSARRLRNARQSRKQRLKADAKKAIQRLQLRTLKQDK	65
QY	269	EIGPDGSCAVCIELKPNDLVRLTCNHFHKTCTVDPMLLKHTKTCMKCKDILKALGIE	328
Db	66	EIGPDGSCAVCIELKPNDLVRLTCNHFHKTCTVDPMLLKHTKTCMKCKDILKALGIE	125
QY	329	VDVEDGSVSLQVPSVNEINSASSHEEDNRSEFASGVSQGTDEPPLEEHVQSTNESL	388
Db	126	VDVEDGSVSLQVPSVNEINSASSHEEDNRSEFASGVSQGTDEPPLEEHVQSTNESL	185
QY	389	QLVNHANSVADVIPHVDNPTFEDETPNQETAVERIKS	428
Db	186	QLVNHANSVADVIPHVDNPTFEDETPNQETAVERIKS	225

RESULT 2
US-09-764-864-1306

```

; Sequence 1306, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT23
; CURRENT APPLICATION NUMBER: US/09/764,864
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1306
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (6)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1306

Query Match      36.8%; Score 838; DB 10; Length 165;
Best Local Similarity 97.5%; Pred. No. 2.3e-76;
Matches 157; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 268 KEIGPDGDSKAVCIYLYKNDLVRLTCNHIHFKTCVDPWLLKHKTCPCMKCDILKALGI 327
Db 5 RXIGPDGDSKAVCIYLYKNDLVRLTCNHIHFKTCVDPWLLKHKTCPCMKCDILKALGI 64

QY 328 EVDVDEGVSQVLPVNSASSHEEDNRSETASSGYASVQGTDEPPLLEHVVOSTNES 387
Db 65 EVDVDEGVSQVLPVNSASSHEEDNRSETASSGYASVQGTDEPPLLEHVVOSTNES 124

QY 388 LQLVNHEANSVAVDVIHYNDNPTFEDETPNQETAVERIKS 428
Db 125 LQLVNHEANSVAVDVIHYNDNPTFEDETPNQETAVERIKS 165

RESULT 3
US-09-935-390A-20
; Sequence 20, Application US/09935390A
; Patent No. US20020076761A1
; GENERAL INFORMATION:
; APPLICANT: Escobedo, Jaime
; Quianjin, Hu
; Garcia, Pablo
; Williams, Lewis T.
; Kothakota, Srinivas
; TITLE OF INVENTION: Secreted Human Proteins
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; Zip: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/935,390A
; FILING DATE: 22-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/988,671
; FILING DATE: 1997-12-11
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane E. R. Potter
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 1369.002

```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 923-2718
; TELEFAX: (510) 655-3542
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 400 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. US20020076761A1e
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-935-390A-20

Query Match      28.9%; Score 656.5; DB 10; Length 400;
Best Local Similarity 37.9%; Pred. No. 1.8e-57;
Matches 161; Conservative 65; Mismatches 148; Indels 51; Gaps 13;

QY 7 AGVSCRGCGFSRLLAFCFLLALSPQAPSGRGAEEAVM-TAYLVNWSRVPHGTGNRTTWEL 65
Db 8 AGVGARG-----VLALA-LLALALCVPGARGALEWFSAVVNIYVDPQT--NLTVMSV 58

QY 66 SEEGVYGQDSPLEPVGVLVPPDPGGA-LNACNPHNTFTVPTVMGSTVQVSWLALIQGG 124
Db 59 SESGRFDSSPKGAHGLGVVPWAPGGDLEGCAPDTRFFVPEP-GRGGAAPWVALVAR-G 116

QY 125 GCTPADKIHLYAYERWASGAVIFNPPGTRNEVIMPSHPGAVDIVAIMIGNLKGTKILQSIQ 184
Db 117 GCTFKDKVLVAARNASAVLYNEERYGNITLPMASHAGTGNIVIMISYKPGREILLELVQ 176

QY 185 RGQVTWVIEVGKHK-GPWVNHYSIFVSVSFFIITAAVTGVYFIFYSGARLRNARQSRK 243
Db 177 KGIPVTMTIGVTRHVOEFISGQSVVFAIAFITMMIISLAWLIFYIYQRFLYTGSQIGS 236

QY 244 QRQLKADAKAIGRLQLRLTKQDKEIGPDGDSKAVCIYLYKNDLVRLTCNHIHFKTC 303
Db- 237 QSHRK-ETKVIQGLLLHTVHKEKGIDVDVAENCVCVCIENFKVKDRIIRILPCKHIFRIC 295

QY 304 VDPWLLKHKTCPCMKCDILKALGI---EVDVDEGVSQVLPVNSASSHEEDNRSE 360
Db 296 IDPWLIDHRTCPMKCLDVIKALGYWGPBGDVQEMPAPESPGRDPAANLSLALPDDGSD 355

QY 361 TASSGYASVQGTDEPPLLEHVVOSTNESLQLVNHEANSVAVDVIHYNDNPTFEDETPNOE 420
Db 356 DSGPPSAS-----PAESEPPQ-----DPSFKGDAGEN-- 382

QY 421 TAVRE 425
Db 383 TALLE 387

RESULT 4
US-09-935-390A-35
; Sequence 35, Application US/09935390A
; Patent No. US20020076761A1
; GENERAL INFORMATION:
; APPLICANT: Escobedo, Jaime
; Quianjin, Hu
; Garcia, Pablo
; Williams, Lewis T.
; Kothakota, Srinivas
; TITLE OF INVENTION: Secreted Human Proteins
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; Zip: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible

```



```

; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1291
; LENGTH: 199
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (187)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1291

Query Match          5.7%; Score 130.5; DB 10; Length 199;
Best Local Similarity 33.8%; Pred. No. 5.5e-05;
Matches 24; Conservative 16; Mismatches 26; Indels 5; Gaps 1;

QY 252 KKAIGRLQLRTLKQGDKEIGPDGSCAVCIELYKPNLVRILTCNHIHFKTCVDPWLLKH 311
Db 125 KEQIDNLAMRSGENDAL-----KTCSCVITEYTEGNKLRKLPSCSHYVHVHCIDRWLSEN 179

QY 312 KTCPCKCKDIL 322
Db 180 STCPICRXAVL 190

RESULT 11
US-09-925-302-638
; Sequence 638, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 638
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-302-638

Query Match          5.7%; Score 130.5; DB 10; Length 207;
Best Local Similarity 33.0%; Pred. No. 5.9e-05;
Matches 29; Conservative 16; Mismatches 32; Indels 11; Gaps 3;

QY 239 AQSRKQRLQADAK----KAIGRLQLRTLKQGDKEIGPDG-----SCAVCIELYKPNL 289
Db 107 SQTRLATQTEEEQRIARIGLIQ--HLPKGVDPGRDGESEKIRECVICMDFVYGP 164

QY 290 VRILTCNHIHFKTCVDPWLLKHKTCPMC 317
Db 165 IRFLPCWHIYHLDCIDDLMRGFTCPSC 192

RESULT 12
US-09-764-864-1288
; Sequence 1288, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

```

```

; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1288
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (83)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (250)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1288

Query Match          5.7%; Score 130; DB 10; Length 461;
Best Local Similarity 28.7%; Pred. No. 0.00023;
Matches 33; Conservative 17; Mismatches 49; Indels 16; Gaps 3;

QY 220 AATVGY----FIFYARRLRNAPQSRKQRLKADA-----KKAIGRLQLRTLK 264
Db 342 AEAISYVDPQLFTYMALEERLAQAMETALAHLESLAVDVEVANPPASKESIDALPEILVT 401

QY 265 QGDKEIGPDGDSCAVCIELYKPNLVRILTCNHIHFKTCVDPWLLKHKTCPMCCKC 319
Db 402 EDHGAVGQE-MCCPICSEYKGEVATELPCHHYFHKPCVSIWLQSGTGPCVCRG 455

RESULT 13
US-09-764-864-832
; Sequence 832, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 832
; LENGTH: 836
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-864-832

Query Match          5.7%; Score 130; DB 10; Length 636;
Best Local Similarity 28.7%; Pred. No. 0.00039;
Matches 33; Conservative 17; Mismatches 49; Indels 16; Gaps 3;

QY 220 AATVGY----FIFYARRLRNAPQSRKQRLKADA-----KKAIGRLQLRTLK 264
Db 517 AEAISYVDPQLFTYMALEERLAQAMETALAHLESLAVDVEVANPPASKESIDALPEILVT 576

QY 265 QGDKEIGPDGDSCAVCIELYKPNLVRILTCNHIHFKTCVDPWLLKHKTCPMCCKC 319
Db 577 EDHGAVGQE-MCCPICSEYKGEVATELPCHHYFHKPCVSIWLQSGTGPCVCRG 630

RESULT 14
US-09-764-864-836
; Sequence 836, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864

```

Search completed: April 22, 2003, 08:34:47
Job time : 43 secs

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 08:24:36 ; Search time 17 Seconds
(without alignments)
740.765 Million cell updates/sec

Title: US-09-854-300-8
Perfect score: 2275
Sequence: 1 MGPPGAGVSCRGCGFSRL.....PTFEDETPNQETAVEREKS 428

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/ptodata/1/iaa/5A COMB pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB pep.*
5: /cgn2_6/ptodata/1/iaa/PTUS COMB pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	468.5	20.6	276	2	US-08-786-606-5
2	331	14.5	284	2	US-08-786-606-9
3	253	11.1	381	2	US-08-867-057-3
4	253	11.1	381	2	US-09-128-369-3
5	246.5	10.8	381	2	US-08-867-057-1
6	246.5	10.8	381	2	US-09-128-369-1
7	152.5	6.7	50	4	US-09-052-089A-15
8	144.5	6.4	40	4	US-09-046-894-44
9	131.5	5.8	104	4	US-09-325-932A-49
10	130	5.7	395	2	US-08-841-349-9
11	127.5	5.6	664	4	US-09-288-140-2
12	114.5	5.0	199	4	US-09-325-932A-46
13	108.5	4.8	551	3	US-08-699-103B-25
14	108.5	4.8	551	4	US-09-229-059-25
15	107.5	4.7	469	2	US-08-968-751-2
16	106.5	4.7	470	4	US-09-052-089A-2
17	103.5	4.5	312	1	US-08-425-061-18
18	103.5	4.5	312	2	US-08-825-886-18
19	103.5	4.5	765	1	US-08-425-061-19
20	103.5	4.5	765	2	US-08-825-886-19
21	103.5	4.5	900	1	US-08-425-061-20
22	103.5	4.5	900	2	US-08-825-886-20
23	103.5	4.5	914	1	US-08-425-061-21
24	103.5	4.5	914	2	US-08-825-886-21
25	103.5	4.5	1202	1	US-08-425-061-22
26	103.5	4.5	1202	2	US-08-825-886-22
27	103.5	4.5	1363	1	US-08-425-061-23

28 103.5 4.5 1363 2 US-08-825-886-23 Sequence 23, Appl
29 103.5 4.5 1852 1 US-08-425-061-24 Sequence 24, Appl
30 103.5 4.5 1852 2 US-08-825-886-24 Sequence 24, Appl
31 103.5 4.5 1863 1 US-08-598-591-2 Sequence 2, Appl
32 103.5 4.5 1863 1 US-08-480-784-2 Sequence 2, Appl
33 103.5 4.5 1863 1 US-08-483-553-2 Sequence 2, Appl
34 103.5 4.5 1863 1 US-08-487-002-2 Sequence 2, Appl
35 103.5 4.5 1863 1 US-08-483-554B-2 Sequence 2, Appl
36 103.5 4.5 1863 1 US-08-798-691-2 Sequence 2, Appl
37 103.5 4.5 1863 1 US-08-798-691-4 Sequence 2, Appl
38 103.5 4.5 1863 1 US-08-798-691-6 Sequence 6, Appl
39 103.5 4.5 1863 1 US-08-488-011B-2 Sequence 2, Appl
40 103.5 4.5 1863 2 US-08-603-753D-2 Sequence 2, Appl
41 103.5 4.5 1863 3 US-08-825-487A-2 Sequence 4, Appl
42 103.5 4.5 1863 3 US-08-825-487A-4 Sequence 4, Appl
43 103.5 4.5 1863 3 US-08-825-487A-6 Sequence 6, Appl
44 103.5 4.5 1863 4 US-09-074-476-2 Sequence 2, Appl
45 103.5 4.5 1863 4 US-09-074-476-4 Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-786-606-5
; Sequence 5, Application US/08786606
; Patent No. 5861495
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Au-Young, Janice
; APPLICANT: Coleman, Roger
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN ZINC-BINDING
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION NUMBER: US/08/786,606
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy RJ
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0173 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 276 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-786-606-5

Query Match 20.6%; Score 468.5; DB 2; Length 276;
Best Local Similarity 52.6%; Pred. No. 9e-44;
Matches 90; Conservative 28; Mismatches 52; Indels 1; Gaps 1;

QY 158 MSHPGAVDIVAIMIGNLKGTKILQSIQGIQVTWIEVCKKHGPM-VNHYISFFVSVSF 216
 Db 1 MTHPGGDIIVAMITELRGKILSVLEKNISVQMTIAVGTMPKPNFSGSLVFSISFI 60
 QY 217 IITAAVGVYFIFYSARLRNARQSRKQRLKADAKAIGRLQLTLKQDKEIGPDGDS 276
 Db 61 VLMISSAWLIFYFKIYTWARNQRLGDAAKKALSKTUTTVKKGDKETDPDPDH 120
 QY 277 CAVCIELYKPNLVRILTCHNHFHKTCDPWLKHKTCPCMKCDILKALGI 327
 Db 121 CAVCIESYKQNDVRIILPCKVFKSCVDPWLSEHCTPCMKCNILKALGI 171

RESULT 2

US-08-786-606-9
 ; Sequence 9, Application US/08786606
 ; Patent No. 5861495
 ; GENERAL INFORMATION:
 ; APPLICANT: Hillman, Jennifer L.
 ; APPLICANT: Au-Young, Janice
 ; APPLICANT: Coleman, Roger
 ; APPLICANT: Goli, Surya K.
 ; TITLE OF INVENTION: NOVEL HUMAN ZINC-BINDING
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: Fast-SEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/786.606
 ; FILING DATE:
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Billings, Lucy RJ
 ; REGISTRATION NUMBER: 36,749
 ; REFERENCE/DOCKET NUMBER: PF-0173 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-855-0555
 ; TELEFAX: 415-845-4166
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 9:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 284 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: GenBank
 ; CLONE: 157535
 ; US-08-786-606-9

Query Match 14.5%; Score 331; DB 2; Length 284;
 Best Local Similarity 34.8%; Pred. No. 2.3e-28;
 Matches 80; Conservative 39; Mismatches 77; Indels 34; Gaps 5;
 QY 165 DIVAIMIGNLKGTKILQSIQGIQVTWIEVCKKHGPMV---NHYISFFVSVFIIITAA 221
 Db 14 NTAAVTYQNIQDLSLTDKGNVTISIEGRGVRITSSILNRTSVLFSISF--IVDD 71
 QY 222 TVGYFIFYSARLRNARQSRKQRLKADAKAIGRLQLTLKQDKEIGPDGDSCAVCI 281

Db 72 ILCWLFYIYQIRFYMQAKDQQRNLCVTKKAIMKIPTKTGKFS-DKDLSDCCAICI 130
 QY 282 ELKYPNDLVRILTCHNHFHKTCDPWLKHKTCPCMKCDILKALG----- 326
 Db 131 EAYKFTDTRILPCKEHFKNCIDPWLIEHRTCPMKCLDLVKFYGVVGDIYQTPSQH 190
 QY 327 -----TEVDVEDGSVSLQVPVSNEISNSASSHE-EDNRSETAS 363
 Db 191 TAPIASTEEVPVIWAVPHGPQLPQLQASNMSSFPAPSHYFQSSRSPSS 240

RESULT 3

US-08-867-057-3
 ; Sequence 3, Application US/08867057
 ; Patent No. 5840535
 ; GENERAL INFORMATION:
 ; APPLICANT: Hillman, Jennifer L.
 ; APPLICANT: Lal, Preeti
 ; APPLICANT: Shah, Purvi
 ; TITLE OF INVENTION: NEW ZINC RING PROTEIN
 ; NUMBER OF SEQUENCES: 3
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/867,057
 ; FILING DATE: Filed Herewith
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Billings, Lucy J.
 ; REGISTRATION NUMBER: 36,749
 ; REFERENCE/DOCKET NUMBER: PF-0311 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-855-0555
 ; TELEFAX: 415-845-4166
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 381 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: GenBank
 ; CLONE: 1321818
 ; US-08-867-057-3

Query Match 11.1%; Score 253; DB 2; Length 381;
 Best Local Similarity 24.0%; Pred. No. 2.1e-19;
 Matches 84; Conservative 60; Mismatches 140; Indels 66; Gaps 14;
 QY 71 YGQDSPLEPVAGVLV--PPDGPALNACNPHNTFTVTWGSTVQVSWIALQGGGCTF 128
 Db 55 FGVLPAEGLKGLFINSKPE-----NACBP----IAPPLRDNSSAFIVLIRR-LECNF 104
 QY 129 ADKHILAYERWASGAVIFNPPGTRNEVIPMSHPG-----AVDIVAIMIGNLKGTKILO-- 181
 Db 105 DIKVLNARAGYKAAIVHNVD--DDLISGNDIEILKIDIPSVFGEASNSLKEEF 162
 QY 182 SIQGIQVTWIEVCKKHGPMVNHYSIFFVSVFIIITAAVTGYFIFYSARLRNARQ 241
 Db 163 TYEKGHWLIPFS-----LPLEVYLPFLIIVGICLILIVIFMITKVFQDRHRA 213

```

QY 242 RKORLAKADAKAIGRLQLRTLKQDKEIGPDGSDCAVCIELYKPNLVRILTCKNHIFHK 301
Db 214 RNRRLKQDLKK-----LPVHKFKGD-----EYDVCAICLDEYDGDGKRLILPCSHAYHC 264
QY 302 TCVDPMWLK-HKTCPCMKCDILKALG-----IEVDVEDGSVSLQVPV-----342
Db 265 KCVDPMWLTKTKTCPCVKQKVVPSQSDSDTSDSQEENEVSNTPLLRPLASVSTQSGF 324
QY 343 -----SNEINSASSHEDNRSETASSGYASVQGTDEPPPLEEHVQSTNE 386
Db 325 ALSSESHQNMTESEYEDNDNDIDSSDAES--GVNEESVVVQLQPNDE 372

RESULT 4
US-09-128-369-3
; Sequence 3, Application US/09128369
; Patent No. 5998372
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: NEW ZINC RING PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/128,369
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/867,057
; FILING DATE: 06/02/1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0311 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 381 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1321818
US-09-128-369-3

```

```

Query Match 11.1%; Score 253; DB 2; Length 381;
Best Local Similarity 24.0%; Pred. No. 2.1e-19;
Matches 84; Conservative 60; Mismatches 140; Indels 66; Gaps 14;

QY 71 YGQDSPLEPVAGLV--PPDGPALNACNPHNTFTVTWGSTVQVSWLALIORGGGCTF 128
Db 55 FGRLPAGLKGFLINSKPE-----NACEP-----IAPPLRDNSSATFIVLIRR-LECNF 104
QY 129 ADKHLAYERWASGAVIFNFGTNEVPMSPHG-----AVDIVAIMGNLKGTKILO-- 181
Db 105 DIKVLNAQRAGYKAAIVHNVD--DDLISMGSDNIDELKIDIPSVFIGASANSKEEF 162
QY 182 SIQRGIQVTWIEVKKHGPWNHYSIFFVSFFIATAVTGYFIFYSARLLENARAQS 241

```

```

Db 163 TYEKGHHVLIPEFS-----LPLEVYLIPFLIIVGICLILIVIFMITKFOODRHA 213
QY 242 RKORLAKADAKAIGRLQLRTLKQDKEIGPDGSDCAVCIELYKPNLVRILTCKNHIFHK 301
Db 214 RNRRLKQDLKK-----LPVHKFKGD-----EYDVCAICLDEYDGDGKRLILPCSHAYHC 264
QY 302 TCVDPMWLK-HKTCPCMKCDILKALG-----IEVDVEDGSVSLQVPV-----342
Db 265 KCVDPMWLTKTKTCPCVKQKVVPSQSDSDTSDSQEENEVSNTPLLRPLASVSTQSGF 324
QY 343 -----SNEINSASSHEDNRSETASSGYASVQGTDEPPPLEEHVQSTNE 386
Db 325 ALSSESHQNMTESEYEDNDNDIDSSDAES--GVNEESVVVQLQPNDE 372

RESULT 5
US-08-867-057-1
; Sequence 1, Application US/08867057
; Patent No. 5840535
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: NEW ZINC RING PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/867,057
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0311 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 381 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BMARNOT02
; CLONE: 104119
US-08-867-057-1

```

```

Query Match 10.8%; Score 246.5; DB 2; Length 381;
Best Local Similarity 24.1%; Pred. No. 1.1e-18;
Matches 83; Conservative 58; Mismatches 132; Indels 71; Gaps 14;

QY 71 YGQDSPLEPVAGLV--PPDGPALNACNPHNTFTVTWGSTVQVSWLALIORGGGCTF 128
Db 55 FGRLPAGLKGFLINSKPE-----NACEP-----IAPPLRDNSSATFIVLIRR-LECNF 104
QY 129 ADKHLAYERWASGAVIFNFGTNEVPMSPHG-----AVDIVAIMGNLKGTKILO-- 181
Db 105 DIKVLNAQRAGYKAAIVHNVD--DDLISMGSDNIDELKIDIPSVFIGESSANSKLDEF 162

```

QY 182 SIQRGIQVMTVEVGGKKGWVNHYSIFFVSVSFFIITAATVGYFIFYSARRLRNARAQS 241
 Db 163 TVEKGGHLLVPEFS-----LPLEYLIPFLIIIVGICLILIVIFWITKVFQDRHRA 213
 QY 242 RKQRQLKADAKAIGRLQLRTLKQDKEIGPDGSCAVCIELYKPNDLVRLITCNHIFPK 301
 Db 214 RRNRRLKQDLKK-----LPVHKFKKGD-----EYDVCAICLDEYEDGDKLRILPCSHAYHC 264
 QY 302 TCVDPMWLK-HKTCPMCKCDILKALG-----IEVDVEDGSVSLQVPV----- 342
 Db 265 KCVDPMWLTKTKTCPCVKQKVPSQSDSDSDTSSQSENEVEHTHTPLLRPLASVSAQSF 324
 QY 343 -----SNEISNSASHEEDNRSETASSGYASVQGTDEPPLEEH 380
 Db 325 ALSESRSHQNTESDYEDDNEEDTDSS-----DAENEINEH 361

RESULT 6

US-09-128-369-1
 ; Sequence 1, Application US/09128369
 ; Patent No. 5998372
 ; GENERAL INFORMATION:
 ; APPLICANT: Hillman, Jennifer L.
 ; APPLICANT: Lal, Preeti
 ; APPLICANT: Shah, Purvi
 ; TITLE OF INVENTION: NEW ZINC RING PROTEIN
 ; NUMBER OF SEQUENCES: 3
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FASTSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/128,369
 ; FILING DATE:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/867,057
 ; FILING DATE: 06/02/1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Billings, Lucy J.
 ; REGISTRATION NUMBER: 36,749
 ; REFERENCE/DOCKET NUMBER: PF-0311 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-855-0555
 ; TELEFAX: 415-845-4166
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 381 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: BMAROT02
 ; CLONE: 104119
 US-09-128-369-1

Query Match 10.8%; Score 246.5; DB 2; Length 381;
 Best Local Similarity 24.1%; Pred. No. 1.1e-18;
 Matches 83; Conservative 58; Mismatches 132; Indels 71; Gaps 14;
 QY 71 YQDPSPLFPVAGLV--PPDGGALNACNPHNTFTVPTVWGSTVQVSWLAIQGGGCTF 128
 Db 55 FGVRLPAGLKGFLINSKPE-----NACEP-----IVPPVKDSSGTFVILIR-LDCNF 104
 QY 129 ADKIHAYERWASGAVINFPPGTRNEVPMHPG-----AVDIVAIMGNLKGFKILQ-- 181

Db 105 DIKVLNQRAGYKAAIVHNVDS--DDLISGNSNDIEVLKKIDIPSVFISSANSLKDEF 162
 QY 182 SIQRGIQVMTVEVGGKKGWVNHYSIFFVSVSFFIITAATVGYFIFYSARRLRNARAQS 241
 Db 163 TVEKGGHLLVPEFS-----LPLEYLIPFLIIIVGICLILIVIFWITKVFQDRHRA 213
 QY 242 RKQRQLKADAKAIGRLQLRTLKQDKEIGPDGSCAVCIELYKPNDLVRLITCNHIFPK 301
 Db 214 RRNRRLKQDLKK-----LPVHKFKKGD-----EYDVCAICLDEYEDGDKLRILPCSHAYHC 264
 QY 302 TCVDPMWLK-HKTCPMCKCDILKALG-----IEVDVEDGSVSLQVPV----- 342
 Db 265 KCVDPMWLTKTKTCPCVKQKVPSQSDSDSDTSSQSENEVEHTHTPLLRPLASVSAQSF 324
 QY 343 -----SNEISNSASHEEDNRSETASSGYASVQGTDEPPLEEH 380
 Db 325 ALSESRSHQNTESDYEDDNEEDTDSS-----DAENEINEH 361

RESULT 7

US-09-052-089A-15
 ; Sequence 15, Application US/09052089A
 ; Patent No. 6346605
 ; GENERAL INFORMATION:
 ; APPLICANT: Lee, Soo Y.
 ; APPLICANT: Choi, Yongwon
 ; TITLE OF INVENTION: SIGNAL TRANSDUCER FOR THE TNF RECEPTOR SUPER
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: David A. Jackson, Esq.
 ; STREET: 411 Hackensack Ave, Continental Plaza, 4th
 ; Floor
 ; CITY: Hackensack
 ; STATE: New Jersey
 ; COUNTRY: USA
 ; ZIP: 07601
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/052,089A
 ; FILING DATE: 31-Mar-1998
 ; CLASSIFICATION: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Jackson Esq., David A.
 ; REGISTRATION NUMBER: 26,742
 ; REFERENCE/DOCKET NUMBER: 600-1-198 CIP 1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 201-487-5800
 ; TELEFAX: 201-343-1684
 ; INFORMATION FOR SEQ ID NO: 15:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 50 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; HYPOTHETICAL: NO
 ; FRAGMENT TYPE: internal
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 15:

US-09-052-089A-15

Query Match 6.7%; Score 152.5; DB 4; Length 50;
 Best Local Similarity 55.6%; Pred. No. 1e-09;
 Matches 25; Conservative 8; Mismatches 11; Indels 1; Gaps 1;
 QY 275 DSCAVCIELYKPNDLVRLITCNHIFHKTCTVDPWLK-HKTCPMCK 318
 Db 2 DVCAICLDEYEDGDKLRILPCSHAYCKVDPWLTKTKTCVPCK 46

```

; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-325-932A-49

Query Match          5.8%; Score 131.5; DB 4; Length 104;
Best Local Similarity 29.4%; Pred. No. 8.5e-07;
Matches 30; Conservative 18; Mismatches 29; Indels 25; Gaps 4;

QY 265 QGDKKEIGPDGSDCAVCIETYKPNLVLRI-TCNHIFHKTCVDPWLLKHKTCPCMCKCDILK 323
Db 17 KGSQKQ----GLECAVCLSKFEDIEILRLPKRHAFTICIDYWLKXHSQCLCR----- 67
QY 324 ALGIEVDVEDGSVSLQVPVNEISN-----SASSHEEDNRSE 360
Db 68 -----HKVSAEDPANFTYNSRMLWSQAMRQDSNLE 99

RESULT 10
US-08-841-349-9
; Sequence 9, Application US/08841349B
; Patent No. 5955594
; GENERAL INFORMATION:
; APPLICANT: MISHRA, LOPA
; TITLE OF INVENTION: GENES CODING PROTEINS FOR EARLY LIVER DEVELOPMENT. . .
; FILE REFERENCE: XX/P04470U50
; CURRENT APPLICATION NUMBER: US/08/841,349B
; CURRENT FILING DATE: 1997-04-30
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Mus musculus
US-08-841-349-9

Query Match          5.7%; Score 130; DB 2; Length 395;
Best Local Similarity 28.7%; Pred. No. 1.3e-05;
Matches 33; Conservative 17; Mismatches 49; Indels 16; Gaps 3;

QY 220 AATGVY----FIFYSARLLRNARQSKRQOLKADA-----KKATGRQLQLRTLK 264
Db 276 AEATSYVDPOFLTMYALEERLAQAMETALAHLESIAVDVEVNPASKESIDALPEILVT 335
QY 265 QGDKKEIGPDGSDCAVCIETYKPNLVLRI-TCNHIFHKTCVDPWLLKHKTCPCMCKC 319
Db 336 EDHGAVGQE-MCCPICCSYKGEVATELPCHHYFHKPCVSIWLQKSGTCPCVCR 389

RESULT 11
US-09-268-140-2
; Sequence 2, Application US/09268140
; Patent No. 6268176
; GENERAL INFORMATION:
; APPLICANT: Gemmill, Robert M.
; APPLICANT: Drabkin, Harry A.
; TITLE OF INVENTION: TRC8, A GENE RELATED TO THE HEDGEHOG RECEPTOR, PATCHED
; FILE REFERENCE: 93445-00004
; CURRENT APPLICATION NUMBER: US/09/268,140
; - CURRENT FILING DATE: 2000-03-12
; PRIOR APPLICATION NUMBER: US 60/077,723
; PRIOR FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 664
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-268-140-2

Query Match          5.6%; Score 127.5; DB 4; Length 664;
Best Local Similarity 25.2%; Pred. No. 6.2e-05;
Matches 38; Conservative 18; Mismatches 44; Indels 51; Gaps 6

```

QY 275 DSCAVCIELYKPNLVRILTCNHIFHKTCVDPWLLKHKTCPMCKCDILKALGIEVDVDEG 334
D 545 DVCAICVHEFTS--ARITPCNHYFHALCKRWLYIQDTCPMCH-----QVYIEDDIKON 598
QY 335 SVSLQVPVNSISNSASSHEEDNRSETASSGYASVQGTDEPPLEHVVQSNESLQVLNHE 394
D 599 S-----NVSN-----NNGFI-----PPNETPEEAVREAAESDRE 628
QY 395 ANSVAVDVIPVDNPTFEEDTNPQETAVRE 425
D 629 LN-----EDDSTDCDDVQRE 644

RESULT 12

US-09-325-932A-46

; Sequence 46, Application US/09325932A
; Patent No. 6451604
; GENERAL INFORMATION:
; APPLICANT: Flinn, Barry
; APPLICANT: Lasham, Annette
; TITLE OF INVENTION: Compositions affecting programmed cell
; TITLE OF INVENTION: death and their use in the modification of forestry plant develop
; FILE REFERENCE: 1022
; CURRENT APPLICATION NUMBER: US/09/325,932A
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 46
; LENGTH: 199
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-325-932A-46

Query Match 5.0%; Score 114.5; DB 4; Length 199;
Best Local Similarity 44.2%; Pred. No. 0.00021;
Matches 19; Conservative 7; Mismatches 16; Indels 1; Gaps 1;

QY 273 DGDSCAVCIELYKPNLVRIL--TCNHIFHKTCVDPWLLKHKTC 314
D 157 DGLECAVCLCEPEENEIARLLPNCRRHNFVCEIDMWFERSHSC 199

RESULT 13

US-08-699-103B-25

; Sequence 25, Application US/08699103B
; Patent No. 6107462
; GENERAL INFORMATION:
; APPLICANT: Rine, Jasper D.
; APPLICANT: Hampton, Randolph
; TITLE OF INVENTION: GENES AND PROTEINS CONTROLLING
; TITLE OF INVENTION: CHOLESTEROL SYNTHESIS
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESS: Fish & Richardson P.C.
; STREET: 2200 Sand Hill Road, Suite 100
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/699,103B
; FILING DATE: 16-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/002,581
; FILING DATE: 17-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Grant D.
; REGISTRATION NUMBER: 31,259

; REFERENCE/DOCKET NUMBER: 09272/005001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/322-5070
; TELEFAX: 650/854-0875
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 551 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: polypeptide
US-08-699-103B-25

Query Match

Best Local Similarity 4.8%; Score 108.5; DB 3; Length 551;
Matches 34; Conservative 26; Mismatches 58; Indels 49; Gaps 6;

QY 273 DGDSCAVCIELYKPNLVRILTCNHIFHKTCVDPWLLKHKTCPMCKCDIL 322
D 345 DDNICIICMDELIHSPNQOTWKNKPKRLPCGHILHLSCLKNMERSQTCPIRLPVF 404
QY 323 KALGIEVDVDEGVSLSQVPVNSISNSASSHEEDNRSETASSGYASVQGTDEPPLEHVVQ 382
D 405 D-----EKNVVQTTFTSNDITTTQTTVDSTGIATDQGF----- 441
QY 383 STNESLQVLNHEANSVAVDVIP-HVDNPTFEEDTNPQETAVREIKS 428
D 442 --NE-----VDLLPRTTSPDIRIVPTQNTIDTLAMRTS 473

RESULT 14

US-09-229-059-25

; Sequence 25, Application US/09229059
; Patent No. 6333172
; GENERAL INFORMATION:
; APPLICANT: Rine, Jasper D.
; APPLICANT: Hampton, Randolph
; TITLE OF INVENTION: GENES AND PROTEINS CONTROLLING
; TITLE OF INVENTION: CHOLESTEROL SYNTHESIS
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESS: Fish & Richardson P.C.
; STREET: 2200 Sand Hill Road, Suite 100
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/229,059
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/699,103
; FILING DATE: 16-AUG-1996
; APPLICATION NUMBER: 60/002,581
; FILING DATE: 17-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Grant D.
; REGISTRATION NUMBER: 31,259
; REFERENCE/DOCKET NUMBER: 09272/005001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/322-5070
; TELEFAX: 650/854-0875
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 551 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: polypeptide
US-09-229-059-25

```

Db 123 QALGKAEMLCSTLKKQKYLEQQOQDETKOQAEARLRS 161
:|: : : : :||| : | :|
Search completed: April 22, 2003, 08:27:49
Job time : 19 secs

```

```

Query Match      4.7%; Score 107.5; DB 2; Length 469;
Best Local Similarity 22.6%; Pred. No. 0.0059;
Matches 36; Conservative 30; Mismatches 82; Indels 11; Gaps 4;

QY 277 CAVCTELYKPNLDVRLITCNHIFIKTCVDPWL--LKHKTCPMC-----KCDILKALGIEV 329
      | | : : : : : | | : : : : : | | : : : : : | | : :
Db 7 CTCISDFPHSDRAAHICGHTFHQLCIOWFETAPSRCTPCQRIQVGKRTIINKLFDEL 66

QY 330 DVEDGSVSLQVPVSNEISNSASSHEEDNRSETASSGVASVQGTDEPFLSEHVQSTNESLQ 389
      | | : : : : : | | : : : : : | | : : : : : | | : :
Db 67 AQEEENLDAEFLKNELDNVAQUSQDKKQRSQ---VVIDLRDTLEER-NATWVSLQ 122

QY 390 LVNHEANSVAVDVIPIHVDNPTFEEDETPNQETAVREIKS 428

```

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 08:16:45 i. Search time 76 Seconds
(without alignments)
750.411 Million cell updates/sec

Title: US-09-854-300-8

Perfect score: 2275

Sequence: 1 MGPPPGAGVSCGGCGFSRL.....PTPEDETNPQETAVREIKS 428

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

```

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	2275	100.0	428	23	Human GRAIL. Homo
2	2255	99.1	428	21	Human protein clon
3	2255	99.1	428	23	ISIGP-4. Homo sap
4	2160	94.9	428	23	Mouse goliath prot
5	2115	93.0	428	23	Synthetic goliath
6	2111	92.8	428	23	Synthetic goliath
7	2111	92.8	428	23	Synthetic goliath
8	2110	92.7	428	23	Synthetic goliath
9	2108	92.7	428	23	Synthetic goliath
10	2106	92.6	428	23	Synthetic goliath

11	2090.5	91.9	427	23	AA047799	Mouse GRAIL. Mus
12	1364	60.0	276	22	AA040136	Human polypeptide
13	1134	49.8	225	22	AAU15897	Human novel secret
14	1027.5	45.2	294	22	AAU25459	Human mdt protein
15	838	36.8	165	22	AAU16353	Human novel secret
16	799	35.1	155	22	AAU1922	Human polypeptide
17	763.5	33.6	305	22	ABG05937	Novel human diagno
18	755.5	33.2	305	21	AA025803	Human secreted pro
19	755.5	33.2	305	22	AA075389	Human secreted pro
20	656.5	28.9	400	19	AA063681	Human secreted pro
21	656.5	28.9	400	22	AA035567	Human polypeptide,
22	650.5	28.6	400	22	AA03452	Human polypeptide,
23	650.5	28.6	400	22	AA088594	Human hydrophobic
24	649.5	28.5	400	23	AB030388	Human polypeptide
25	617.5	27.1	419	22	AA088558	Human hydrophobic
26	617.5	27.1	419	23	AB077003	Human protein sequ
27	617.5	27.1	419	23	AB077025	Human protein sequ
28	617.5	27.1	419	23	AB064937	Human albumin fusi
29	617.5	27.1	419	23	AB064938	Human albumin fusi
30	617.5	27.1	419	23	AB030389	Human polypeptide
31	592.5	26.0	419	23	AAU74921	Mouse g1 protein s
32	586	25.8	462	23	AAU74919	Human goliath prot
33	551	24.2	115	23	AB030387	Human polypeptide
34	545	24.0	341	22	AB051081	Human secreted pro
35	537	23.6	312	22	AB011857	Human secreted pro
36	537	23.6	341	20	AA079358	Human protein SEQ
37	537	23.6	341	20	AA089003	Polypeptide fragme
38	532.5	23.4	210	21	AA041793	Human ORFX ORF1557
39	525	23.1	257	22	AB012676	Novel human diagno
40	468.5	20.6	276	19	AA069602	Human zinc binding
41	468.5	20.6	276	19	AA036966	Human secreted pro
42	468.5	20.6	276	23	AAU74920	Human goliath-2 pr
43	462.5	20.3	276	20	AA088996	Polypeptide fragme
44	462.5	20.3	276	22	AB051074	Human secreted pro
45	438.5	19.3	167	20	AA088995	Polypeptide fragme

ALIGNMENTS

RESULT 1	AA047800	
ID	AA047800 standard; Protein; 428 AA.	
XX	AA047800;	
AC	AA047800;	
XX		
DT	25-MAR-2002 (first entry)	
XX		
DE	Human GRAIL.	
XX		
KW	Murine; human; GRAIL; energy; attenuation; tyrosine; phosphorylation; antigenic stimulation; interleukin-2; gene therapy; polymorphism; IL-2; autoimmune disease; tumour cell; cancer; transplant rejection.	
XX		
OS	Homo sapiens.	
XX		
PH	Key	Location/Qualifiers
FT	Domain	339..413
FT		/label= Zinc_RING_finger_domain
FT		/note= "Mediates protein/protein interaction"
XX		
PN	WO200185943-A1.	
XX		
PD	15-NOV-2001.	
XX		
PF	11-MAY-2001; 2001WO-US15385.	
XX		
PR	11-MAY-2000; 2000US-203513P.	
XX		
PA	(STFD) UNIV LELAND STANFORD JUNIOR.	
XX		
PI	Ford GS, Bloom D, Fathman CG;	
XX		

DR WPI; 2002-055597/07.
DR N-PSDB; AA172095.
XX Novel nucleic acid sequences of energy associated genes, including
PT GRAIL gene useful in the evaluation of pathophysiology or immunotherapy
PT of cancer, autoimmune disease and transplant rejection -
XX Claim 13; Page 47-48; 50pp; English.
PS This sequence shows human GRAIL protein. The GRAIL gene is an
XX energy associated gene which is upregulated during the early stages
CC of induction of energy. GRAIL has been shown to attenuate IL-2
CC transcription in T-cells during response to antigenic stimulation.
CC GRAIL DNA is useful for decreasing the responsiveness of a T cell
CC population, especially synthesis of interleukin-2 (IL-2) in response to
CC antigenic stimulation, by up-regulating GRAIL activity in the T cell
CC population. GRAIL DNA is useful in producing compositions that modulate
CC induction or maintenance of energy, for gene therapy, mapping
CC functional regions of the encoded protein, to analyse a patient sample
CC for the presence of polymorphisms or alterations in expression of
CC sequences associated with T cell energy, disease states, genetic
CC predisposition to a disease state, and in studying associated
CC physiological pathways. Modulation of the gene activity in vivo is
CC useful for prophylactic and therapeutic purposes, such as treating
CC autoimmune disease and to enhance immune response to tumour cells, and
CC identification of energetic T cells. GRAIL DNA is useful in the
CC evaluation of the pathophysiology or immunotherapy of cancer,
CC autoimmune disease, and transplant rejection. Genetic sequences
CC involved in energy induction are useful as markers in the evaluation of
CC specific immunotherapies. Functional characterization of genes involved
CC in energy induction allows the elucidation of the mechanisms of T cell
CC energy, including the transcriptional blockade of IL-2, which may be
CC manipulated to regulate T cell responses in human disease. Modulation
CC of expression of GRAIL gene is useful in manipulating the anergic
CC state. The genetic sequences find use alone or in combinations in
CC determining the expression profile of cells relating to energy, for
CC example in screening of candidate biologically active compounds for
CC modulation of T cell energy. Expression of GRAIL has been found to
CC block tyrosine phosphorylation of a protein present in T cells during
XX activation.
SQ Sequence 428 AA;
Query Match 100.0%; Score 2275; DB 23; Length 428;
Best Local Similarity 100.0%; Pred. No. 1.3e-236;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGPPPGAGVSCRGCGFRLAWCFLLALSPQAPGSGAEAVWTAYLNVSRVPHTVGNR 60
DB 1 MGPPPGAGVSCRGCGFRLAWCFLLALSPQAPGSGAEAVWTAYLNVSRVPHTVGNR 60
QY 61 TWELSEGVGVQDSPLFVAGLVPPDGPALNACNPHNTFTVPTWGSTVQVSWLALI 120
DB 61 TWELSEGVGVQDSPLFVAGLVPPDGPALNACNPHNTFTVPTWGSTVQVSWLALI 120
QY 121 QRGCGCTFADKIHILAYERWASGAVIFNPPGTRNEVIMPSHPGAVDVAIMIGNLKGTKIL 180
DB 121 QRGCGCTFADKIHILAYERWASGAVIFNPPGTRNEVIMPSHPGAVDVAIMIGNLKGTKIL 180
QY 181 QSIQRIQVTVWIEVKGKGPWNHYSIFFVSVSFFIITAATVGYFFYSARRLRNARAQ 240
DB 181 QSIQRIQVTVWIEVKGKGPWNHYSIFFVSVSFFIITAATVGYFFYSARRLRNARAQ 240
QY 241 SRKQRLKADAKKAIGRLQLRTLKQDKIEIGPDGSCAVCIYKPKNDLVRILTCNHFH 300
DB 241 SRKQRLKADAKKAIGRLQLRTLKQDKIEIGPDGSCAVCIYKPKNDLVRILTCNHFH 300
QY 301 KTCVDPWLLKHKTCPMKCKDILKALGIEVDVDSVSLQVPVSNISASSHEEDNRSE 360
DB 301 KTCVDPWLLKHKTCPMKCKDILKALGIEVDVDSVSLQVPVSNISASSHEEDNRSE 360
QY 361 TASSGVVQSGTDEPPEHVSQTNESLQVNVNHEANSVAVDVIPHVDNPTFEDETPE 420
DB 361 TASSGVVQSGTDEPPEHVSQTNESLQVNVNHEANSVAVDVIPHVDNPTFEDETPE 420

DB 361 TASSGVVQSGTDEPPEHVSQTNESLQVNVNHEANSVAVDVIPHVDNPTFEDETPE 420
QY 421 TAVREIKS 428
DB 421 TAVREIKS 428
RESULT 2
AAV94897
ID AAV94897 standard; Protein; 428 AA.
AC AAV94897;
XX 12-JUN-2000 (first entry)
DE Human protein clone HP10574.
XX Human protein; hydrophobic domain; nutritional source; haematopoiesis;
KW cytokine production; cell proliferation; cell differentiation;
KW immune deficiency; infectious disease; autoimmune disorder; asthma;
KW multiple sclerosis; systemic lupus erythematosus; rheumatoid arthritis;
KW allergic reaction; osteoporosis; osteoarthritis; periodontal disease;
KW nervous system disorder; Alzheimer's disease; Parkinson's disease;
KW Huntington's disease; liver fibrosis; lung fibrosis; reperfusion injury;
KW systemic cytokine damage; tissue differentiation; inflammatory condition;
KW coagulation disorder; myocardial infarction; septic shock; stroke;
KW septic shock; sepsis; ischaemia; reperfusion injury; arthritis; tumour;
KW nephritis; therapy.
XX Homo sapiens.
OS WO200005367-A2.
PN 03-FEB-2000.
PD 22-JUL-1999; 99WO-JP03929.
XX 24-JUL-1998; 98JP-0208820.
PR 07-AUG-1998; 98JP-0224105.
PR 25-AUG-1998; 98JP-0238116.
PR 09-SEP-1998; 98JP-0254736.
PR 29-SEP-1998; 98JP-0275505.
XX (SAGA) SAGAMI CHEM RES CENT.
PA (PROT-) PROTEGENE INC.
XX Kato S, Kimura T;
XX WPI; 2000-182694/16.
XX Novel human proteins having hydrophobic domains useful for treating
PT osteoporosis, Alzheimer's disease, Parkinson's disease, asthma,
PT multiple sclerosis, rheumatoid arthritis, cancer, anaemia, and stroke -
XX Claim 1; Page 323-324; 351pp; English.
XX This sequence represents a human protein of the invention, which has
CC hydrophobic domains. The DNA sequences can be used as a probe or as a
CC genetic marker. The protein can also be used as a marker, and to identify
CC potential genetic disorders. The DNA and protein can also be used as
CC nutritional sources or supplements. The protein exhibits cytokine, cell
CC proliferation, cell differentiation activities and induces production of
CC other cytokines in certain cell populations. The protein also exhibits
CC immune stimulating or immune suppressing activity. It can be used in the
CC treatment of various immune deficiencies and disorders, and to treat
CC infectious diseases caused by viral, bacterial, fungal or other
CC infections. The protein is also used for treating autoimmune disorders
CC such as multiple sclerosis, systemic lupus erythematosus, and rheumatoid
CC arthritis. It is also useful in the treatment of allergic reactions and
CC conditions such as asthma, and in immune suppression after organ
CC transplantation. The protein is useful in regulation of haematopoiesis
CC and consequently in the treatment of myeloid or lymphoid cell
CC deficiencies. It is also used in compositions for tissue growth or

CC regeneration. The protein is also used in the treatment of osteoporosis
 CC or osteoarthritis and in the treatment of periodontal disease and other
 CC tooth repair processes. The protein is used in the treatment of nervous
 CC system disorders such as Alzheimer's disease, Parkinson's disease, and
 CC Huntington's disease. They are useful for protection or regeneration and
 CC treatment of lung or liver fibrosis, reperfusion injury in various
 CC tissues, and conditions resulting from systemic cytokine damage. They are
 CC also used for promoting or inhibiting tissue differentiation. They are
 CC also used as contraceptives since they exhibit activin or inhibin related
 CC activities and as a fertility inducing therapeutic. They are used for
 CC treating various coagulation disorders and in treatment and prevention of
 CC conditions resulting from coagulation activities e.g. myocardial
 CC infarction or stroke. They also acts as receptors, receptor ligands or
 CC inhibitors or agonists of receptor/ligand interactions. They are used to
 CC treat inflammatory conditions such as septic shock, sepsis, ischaemia
 CC reperfusion injury, arthritis, and nephritis. They can be used to
 CC prevent tumours.
 XX
 SQ Sequence 428 AA;

Query Match 99.1%; Score 2255; DB 21; Length 428;
 Best Local Similarity 99.3%; Pred. No. 1.9e-234;
 Matches 425; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGPPPGAGVSCRGCGSRLLAWCFLLALSPQAPGSRGAENVMTAYLVNSWRVPHTGVNR 60
 DB 1 MGPPPGAGVSCRGCGSRLLAWCFLLALSPQAPGSRGAENVMTAYLVNSWRVPHTGVNR 60
 QY 61 TWELSEGVYQDSPLPVPVAGVLPDPGPGALNACNPHNTFTVPTVMGSTVQVSWLALI 120
 DB 61 TWELSEGVYQDSPLPVPVAGVLPDPGPGALNACNPHNTFTVPTVMGSTVQVSWLALI 120
 QY 121 QRGCGCTFADKIHILAYERWASGAVIFNPGTRNEVPMSPHGAVIDIVAIMGNLKGTKIL 180
 DB 121 QRGCGCTFADKIHILAYERWASGAVIFNPGTRNEVPMSPHGAVIDIVAIMGNLKGTKIL 180
 QY 181 QSIQRGQVTWVIEVGKKGHPWNHYSIFFVSVSPFFITTAATVGYFTFYSGARLRNARQ 240
 DB 181 QSIQRGQVTWVIEVGKKGHPWNHYSIFFVSVSPFFITTAATVGYFTFYSGARLRNARQ 240
 QY 241 SRQRQLKADAKAIGRLQRLTLKQDKETGPDGDSCAVCIYKPNDLVRLITCNHIFH 300
 DB 241 SRQRQLKADAKAIGRLQRLTLKQDKETGPDGDSCAVCIYKPNDLVRLITCNHIFH 300
 QY 301 KTCVDPMLLKHKTCPMKCKDILKALGIEVDVDSVLSQVPVSNISNSASSHEEDNRSE 360
 DB 301 KTCVDPMLLKHKTCPMKCKDILKALGIEVDVDSVLSQVPVSNISNSASSHEEDNRSE 360
 QY 361 TASSGYASVQGTDEPPPLEEHVQSTNESLQLVNHEANSVAVDVIPHDVNPFTFEDETPNQE 420
 DB 361 TASSGYASVQGTDEPPPLEEHVQSTNESLQLVNHEANSVAVDVIPHDVNPFTFEDETPNQE 420
 QY 421 TAVREIKS 428
 DB 421 TAVREIKS 428

RESULT 3

AAB47874
 ID AAB47874 standard; Protein; 428 AA.

XX AC AAB47874;

XX DT 15-APR-2002 (first entry)

XX DE ISIGP-4.

XX Human; intracellular signalling protein; ISIGP; gene; autoimmune;
 KW cell proliferation; inflammation; gastrointestinal disorder;
 KW reproductive disorder; developmental disorder.

XX OS Homo sapiens.

XX DR

Key	Location/Qualifiers
Key Peptide	1..41
FT	/label= Signal peptide
FT	/note= "identified using HMW and SPSCAN"
FT	10
FT	/note= "potentially O-phosphorylated"
FT	48
FT	/note= "potentially N-glycosylated"
FT	50
FT	/note= "potentially O-phosphorylated"
FT	59
FT	/note= "potentially N-glycosylated"
FT	61
FT	/note= "potentially O-phosphorylated"
FT	75
FT	/note= "potentially O-phosphorylated"
FT	101
FT	/note= "potentially N-glycosylated"
FT	127
FT	/note= "potentially O-phosphorylated"
FT	151
FT	/note= "potentially O-phosphorylated"
FT	206..229
FT	/label= Transmembrane domain
FT	/note= "identified using HMW"
FT	230
FT	/note= "potentially O-phosphorylated"
FT	241
FT	/note= "potentially O-phosphorylated"
FT	262
FT	/note= "potentially O-phosphorylated"
FT	273..323
FT	/label= Zinc finger_C3CH4 type
FT	/note= "identified using BLAST_DOMO"
FT	273..328
FT	/label= Zinc finger_C3CH4 type
FT	/note= "identified using PROFILESSCAN"
FT	277..317
FT	/label= Zinc finger_C3CH4 type
FT	/note= "identified using HMW_PPFAM"
FT	292..306
FT	/label= PHD finger
FT	/note= "identified using BLIMPS_PPFAM"
FT	302
FT	/note= "potentially O-phosphorylated"
FT	351
FT	/note= "potentially O-phosphorylated"
FT	352
FT	/note= "potentially O-phosphorylated"
FT	357
FT	/note= "potentially N-glycosylated"
FT	383
FT	/note= "potentially O-phosphorylated"
FT	385
FT	/note= "potentially N-glycosylated"
FT	410
FT	/note= "potentially O-phosphorylated"
XX	WO200194391-A2.
PN	13-DEC-2001.
XX	07-JUN-2001; 2001WO-US18595.
XX	08-JUN-2000; 2000US-210582P.
PR	16-JUN-2000; 2000US-212443P.
XX	(INCY-) INCYTE GENOMICS INC.
PA	Yue H, He A, Nguyen DB, Yao MG, Bandman O, Burford N, Tang YT;
PI	Xu Y, Hafalia A, Azimzai Y, Walia NK;
XX	WPI; 2002-154564/20.

DR N-PSDB; AAU72321.

XX New human intracellular signaling protein and polynucleotides useful
 PT for diagnosing, treating or preventing cell proliferative,
 PT autoimmune/inflammatory, gastrointestinal, reproductive and
 PT developmental disorders -

XX Claim 1; Page 99-100; 106pp; English.

XX The sequences given in ABA47971-75 show novel human intracellular
 CC signalling proteins (ISIGP). The polynucleotides and ISIGP proteins
 CC may be used for the diagnosis, treatment or prevention of cell
 CC proliferative, autoimmune/inflammatory, gastrointestinal, reproductive
 CC and developmental disorders. This protein has homology to the mouse
 CC gl-related zinc finger protein.

XX Sequence 428 AA;

Query Match 99.1%; Score 2255; DB 23; Length 428;
 Best Local Similarity 99.3%; Pred. No. 1.9e-234;
 Matches 425; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MGPPPGAGVSCRGCGFSRLAWCFLLALSPQAPGSRGAENVMTAYLNVSWRPVHTGVNR 60
 Db 1 MGPPPGAGVSCRGCGFSRLAWCFLLALSPQAPGSRGAENVMTAYLNVSWRPVHTGVNR 60

Qy 61 TWELSEGVYQDPSLEPVAGVLVPPDGPALNACNPHNTFTVPTVMGSTVQVSWLALI 120
 Db 61 TWELSEGVYQDPSLEPVAGVLVPPDGPALNACNPHNTFTVPTVMGSTVQVSWLALI 120

Qy 121 QRGGGCTFADKIHILAYERWASGAVIFNPPGTRNEVIMPSHPGAVDIIVAIMIGNLKGTKIL 180
 Db 121 QRGGGCTFADKIHILAYERWASGAVIFNPPGTRNEVIMPSHPGAVDIIVAIMIGNLKGTKIL 180

Qy 181 QSIORGIOVTWIEVGKKGPMWVNHYSIFFVSVPFIITTAATVGYPIFYSARLRNARAQ 240
 Db 181 QSIORGIOVTWIEVGKKGPMWVNHYSIFFVSVPFIITTAATVGYPIFYSARLRNARAQ 240

Qy 241 SRKQRLKADAKAIGRLQLRTLKQDKEIGPDGSCAVCIELYKPNDLVRLTNCNHFH 300
 Db 241 SRKQRLKADAKAIGRLQLRTLKQDKEIGPDGSCAVCIELYKPNDLVRLTNCNHFH 300

Qy 301 KTCVDPWLLKHTCPMKCDILKALGIEVDVDSVLSQVPSNEISNSASHEEDNRSE 360
 Db 301 KTCVDPWLLKHTCPMKCDILKALGIEVDVDSVLSQVPSNEISNSASHEEDNRSE 360

Qy 361 TASSGYASVQGTDEPPLLEHVQSTNESLQVNHANSVAVDVI PHVDNPTFEDETNPQE 420
 Db 361 TASSGYASVQGTDEPPLLEHVQSTNESLQVNHANSVAVDVI PHVDNPTFEDETNPQE 420

Qy 421 TAVREIKS 428
 Db 421 TAVREIKS 428

RESULT 4
 AAU74918
 ID AAU74918 standard; Protein; 428 AA.

XX AAU74918;
 XX 23-APR-2002 (first entry)

DE Mouse goliath protein sequence.

XX Mouse; goliath protein; antiangiogenic; vasotropic; gene therapy;
 KW dosage form; angiogenesis; neurogenesis; tumour; vascularisation;
 KW cancer; ischaemia; neuroblastoma; neurodegenerative disease;
 KW unrecovered nerve trauma.

XX Mus sp.
 XX Key Location/Qualifiers

FT Peptide 20..32
 FT Protein /label= Signal_sequence
 FT Domain /label= Mature_goliath_protein
 FT Domain /label= Extracellular_N_terminal_domain
 FT Domain /label= Transmembrane_domain
 FT Domain /label= Intracellular_C_terminal_domain
 FT Domain /label= C3H2C3_ring_finger_motif

PN WO200193681-A1.
 XX 13-DEC-2001.
 XX 01-JUN-2001; 2001WO-US18000.
 XX 02-JUN-2000; 2000US-0586398.
 XX (RECC) UNIV CALIFORNIA.
 PA Harland R, Baker JC;
 XX WPI; 2002-147637/19.
 XX N-PSDB; ASK12983.
 DR New compositions comprising goliath proteins, useful for modulating
 XX angiogenesis or neurogenesis in mammals e.g. for preventing or
 XX treating undesirable vascularisation of a tumour, ischaemia or
 XX neurodegenerative disease -
 PS Claim 1; Page 29-30; 45pp; English.

XX The present invention relates to a new pharmaceutical composition that
 CC comprises a goliath polypeptide in dosage form. The goliath polypeptide
 CC has a sequence identity of at least 75% to the protein sequences
 CC (AAU74918-AAU74921) fully defined in the specification. The composition
 CC is useful for modulating angiogenesis or neurogenesis in mammals,
 CC particularly in humans or mice. Specifically, the composition is useful
 CC for the prophylactic and/or therapeutic treatment of excess angiogenesis
 CC e.g. undesirable vascularisation of a tumour or insufficient angiogenesis
 CC e.g. ischaemia, excess neurogenesis e.g. neuroblastoma or insufficient
 CC neurogenesis e.g. neurodegenerative disease or unrecovered nerve trauma.
 CC The present amino acid sequence represents the mouse goliath protein
 CC that was used in the invention for modulating angiogenesis or
 CC neurogenesis.

SQ Sequence 428 AA;

Query Match 94.9%; Score 2160; DB 23; Length 428;
 Best Local Similarity 94.6%; Pred. No. 3.5e-224;
 Matches 405; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

Qy 1 MGPPPGAGVSCRGCGFSRLAWCFLLALSPQAPGSRGAENVMTAYLNVSWRPVHTGVNR 60
 Db 1 MGPPPGAGVSCRGCGFSRLAWCFLLALSPQAPGSRGAENVMTAYLNVSWRPVHTGVNR 60

Qy 61 TWELSEGVYQDPSLEPVAGVLVPPDGPALNACNPHNTFTVPTVMGSTVQVSWLALI 120
 Db 61 TWELSEGVYQDPSLEPVAGVLVPPDGPALNACNPHNTFTVPTVMGSTVQVSWLALI 120

Qy 121 QRGGGCTFADKIHILAYERWASGAVIFNPPGTRNEVIMPSHPGAVDIIVAIMIGNLKGTKIL 180
 Db 121 QRGGGCTFADKIHILAYERWASGAVIFNPPGTRNEVIMPSHPGAVDIIVAIMIGNLKGTKIL 180

Qy 181 QSIORGIOVTWIEVGKKGPMWVNHYSIFFVSVPFIITTAATVGYPIFYSARLRNARAQ 240
 Db 181 QSIORGIOVTWIEVGKKGPMWVNHYSIFFVSVPFIITTAATVGYPIFYSARLRNARAQ 240

Qy 241 SRKQRLKADAKAIGRLQLRTLKQDKEIGPDGSCAVCIELYKPNDLVRLTNCNHFH 300
 Db 241 SRKQRLKADAKAIGRLQLRTLKQDKEIGPDGSCAVCIELYKPNDLVRLTNCNHFH 300

Db 241 SRKORQLKADAKKAIGKQLRLTLKQDKKEIGPDGDSCAVCIELYKPNDLVRLITCNHIFH 300
 Qy 301 KTCVDPMLLKHTCTPMCKCDILKALGIEVDGVSLSQVPVSNEISNSASSHEEDNRSE 360
 Db 301 KTCVDPMLLEHRTCTPMCKCDILKALGIEVDGVSLSQVPVSNEASNTASPHEDSRSE 360
 Qy 361 TASSGYASVQGTDEPPLEEHVQSTNESLQLVNHEANSVADVIPHVDNPTFEEDTPNOE 420
 Db 361 TASSGYASVQGADEPPLEEHQAQSANENLQLVNHEANSVADVIPHVDNPTFEEDTPDOE 420
 Qy 421 TAVREIKS 428
 Db 421 AAVREIKS 428

RESULT 5
 AAU74927
 ID AAU74927 standard; Protein; 428 AA.
 XX
 AC AAU74927;
 XX
 XX 23-APR-2002 (first entry)
 DT
 DE
 DE
 XX
 XX Synthetic gliath protein sequence #6.
 KW Gliath protein; antiangiogenic; vasotropic; gene therapy;
 KW dosage form; angiogenesis; neurogenesis; tumour; vascularisation;
 KW cancer; ischaemia; neuroblastoma; neurodegenerative disease;
 KW unrecovered nerve trauma.
 XX
 OS Synthetic.
 XX
 XX WO200193681-A1.
 PN
 PD 13-DEC-2001.
 XX
 XX 01-JUN-2001; 2001WO-US18000.
 XX
 XX 02-JUN-2000; 2000US-0586398.
 PR
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 XX Harland R, Baker JC;
 XX
 XX WPI; 2002-147637/19.
 XX
 XX New compositions comprising gliath proteins, useful for modulating
 PT angiogenesis or neurogenesis in mammals e.g. for preventing or
 PT treating undesirable vascularisation of a tumour, ischaemia or
 PT neurodegenerative disease -
 XX
 XX Disclosure; Page 37-38; 45pp; English.
 PS
 XX
 XX The present invention relates to a new pharmaceutical composition that
 CC comprises a gliath polypeptide in dosage form. The gliath polypeptide
 CC has a sequence identity of at least 75% to the protein sequences
 CC (AAU74918-AAU74921) fully defined in the specification. The composition
 CC is useful for modulating angiogenesis or neurogenesis in mammals,
 CC particularly in humans or mice. Specifically, the composition is useful
 CC for the prophylactic and/or therapeutic treatment of excess angiogenesis
 CC e.g. undesirable vascularisation of a tumour or insufficient angiogenesis
 CC e.g. ischaemia, excess neurogenesis e.g. neuroblastoma or insufficient
 CC neurogenesis e.g. neurodegenerative disease or unrecovered nerve trauma.
 CC The present amino acid sequence represents synthetic gliath protein #6
 CC that is one of several artificial gliath proteins (AAU74922-AAU74927)
 CC used in the invention for modulating angiogenesis or neurogenesis.
 XX
 SQ Sequence 428 AA;

Query Match 93.0%; Score 215; DB 23; Length 428;
 Best Local Similarity 92.8%; Pred. No. 2.6e-219;
 Matches 397; Conservative 11; Mismatches 20; Indels 0; Gaps 0;

Qy 1 MGPPPGAGVSCRGCGFSRLLAWCFLLALSPQAPGSRGAENVMTAYLNVSRVPHTGVNR 60
 Db 1 MGPPPGIGVYCRGCGGAARLLAWCFLLALSPHAPGSRGAENVMTAYLNVGRVPHHTGVNR 60
 Qy 61 TWELSEBEGVYGQSPLEPVPVAGVLVPPDGPALNACNPHNTFTVPTVMGSTVQVSWLALI 120
 Db 61 TWELSEBEGVYGQSPLEPVPVAGVLVPPDGPALNACNPHNTFTVPTVMGSTVQVSWLALI 120
 Qy 121 QRGCGCTPADKIHILAYERWASGAVIFNPPGTRNEVIMSHPGAVDIIVAIMIGNLKGTKIL 180
 Db 121 QRGCGCTPADKIHILASERGASGAVIFNPPGIRNEVIMSHPGAGDIIVAIMIGNLKGTKIL 180
 Qy 181 OSIORGIOVTWVIEVGKKGPMVNHYSITFFVSVSFFIITATVGYFIFYSARRLRNARAO 240
 Db 181 OSIORGIOVTWVIEVGKKGPMVNHYSITFFVSVSFFIITATVGYFIFYSARRLRNARAO 240
 Qy 241 SRKORQLKADAKKAIGKQLRLTLKQDKKEIGPDGDSCAVCIELYKPNDLVRLITCNHIFH 300
 Db 241 SRKORQLKADLKAIGKQLRLTLKQDKKEIGPDGDSCAVCIELYKPNDLVRLITCNHIFH 300
 Qy 301 KTCVDPMLLKHTCTPMCKCDILKALGIEVDGVSLSQVPVSNEISNSASSHEEDNRSE 360
 Db 301 MTCVDPMLLEHRTCTPMCKCDILKALGIEVDGVSLSQVPVSNEASNTANPHEDSRSE 360
 Qy 361 TASSGYASVQGTDEPPLEEHVQSTNESLQLVNHEANSVADVIPHVDNPTFEEDTPNOE 420
 Db 361 TASSGYASVQGADEPPLEEHQAQSANENLQLVNHEANSVADVIPHVDNPTFEEDTPDOE 420
 Qy 421 TAVREIKS 428
 Db 421 AAVREIKS 428

RESULT 6
 AAU74922
 ID AAU74922 standard; Protein; 428 AA.
 XX
 AC AAU74922;
 XX
 XX 23-APR-2002 (first entry)
 DT
 DE
 DE
 XX
 XX Synthetic gliath protein sequence #1.
 KW Gliath protein; antiangiogenic; vasotropic; gene therapy;
 KW dosage form; angiogenesis; neurogenesis; tumour; vascularisation;
 KW cancer; ischaemia; neuroblastoma; neurodegenerative disease;
 KW unrecovered nerve trauma.
 XX
 OS Synthetic.
 XX
 XX WO200193681-A1.
 PN
 PD 13-DEC-2001.
 XX
 XX 01-JUN-2001; 2001WO-US18000.
 XX
 XX 02-JUN-2000; 2000US-0586398.
 PR
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 XX Harland R, Baker JC;
 XX
 XX WPI; 2002-147637/19.
 XX
 XX New compositions comprising gliath proteins, useful for modulating
 PT angiogenesis or neurogenesis in mammals e.g. for preventing or
 PT treating undesirable vascularisation of a tumour, ischaemia or
 PT neurodegenerative disease -
 XX
 XX Disclosure; Page 32-33; 45pp; English.
 PS
 XX
 XX The present invention relates to a new pharmaceutical composition that
 CC comprises a gliath polypeptide in dosage form. The gliath polypeptide

CC has a sequence identity of at least 75% to the protein sequences
 CC (AAU74921-AAU74921) fully defined in the specification. The composition
 CC is useful for modulating angiogenesis or neurogenesis in mammals,
 CC particularly in humans or mice. Specifically, the composition is useful
 CC for the prophylactic and/or therapeutic treatment of excess angiogenesis
 CC e.g. undesirable vascularisation of a tumour or insufficient angiogenesis
 CC e.g. ischaemia, excess neurogenesis e.g. neuroblastoma or insufficient
 CC neurogenesis e.g. neurodegenerative disease or unrecovered nerve trauma.
 CC The present amino acid sequence represents synthetic goliath protein #1
 CC that is one of several artificial goliath proteins (AAU74922-AAU74927)
 CC used in the invention for modulating angiogenesis or neurogenesis.

XX Sequence 428 AA;

Query Match 92.8%; Score 2111; DB 23; Length 428;
 Best Local Similarity 92.8%; Pred. No. 6.9e-219;
 Matches 397; Conservative 10; Mismatches 21; Indels 0; Gaps 0;

QY 1 MGPPPGAGVSCRGCGFRRLLAWCFLLALSPAPGSRGAEAVWTAYLNVSRVPHGTGVR 60
 DB 1 MGPPPGIGVYCRGGCGAARLLAWCFLLALSPAPGSRGAEAVWTAYLNVSRVPHGTGVR 60

QY 61 TWELSEGVYQDSPLPVPVAGLVPPDPGALNACNPHNTFTVPTVMGSTVQVSWLALI 120
 DB 61 TWELSEGVYQDSPLPVPVAGLVPPDPGALNACNPHNTFTVPTVMGSTVQVSWLALI 120

QY 121 QRGGGCTPADKIHILASERWASGAVIFNPPGTRNEVPMHSHPGAVDIVAIMIGNLKGTKIL 180
 DB 121 QRGGGCTPADKIHILASERWASGAVIFNPPGTRNEVPMHSHPGAVDIVAIMIGNLKGTKIL 180

QY 181 QSIQRIQVTWVIEVKGKHGPNWVNHYSIFFVSVPFFITTAATVGYFIFYSARRLRNARAQ 240
 DB 181 QSIQRIQVTWVIEVKGKHGPNWVNHYSIFFVSVPFFITTAATVGYFIFYSARRLRNARAQ 240

QY 241 SRKQRLKADAKKATIGRLQLRTLKQDKEIGPDGDSCAVCIELYKPNDLVRLITCNHIFH 300
 DB 241 SRKQRLKADAKKATIGRLQLRTLKQDKEIGPDGDSCAVCIELYKPNDLVRLITCNHIFH 300

QY 301 KTCVDPWLLKHKTCPMKCKDILKALGIEVDVEDGVSLSQVPSVNEISNSASSHEEDNRSE 360
 DB 301 NTCVDPWLLKHKTCPMKCKDILKALGIEVDVEDGVSLSQVPSVNEISNSASSHEEDNRSE 360

QY 361 TASSGYASVOGTDEPPELHHVQSTNESLQVNVHANSVAVDVIVPHVDNPTFEEDTPOE 420
 DB 361 TASSGYASVOGADEPPELHHVQSTNESLQVNVHANSVAVDVIVPHVDNPTFEEDTPOE 420

QY 421 TAVREIKS 428
 DB 421 AAVREIKS 428

RESULT 7
 ID AAU74925
 AC AAU74925;
 XX AAU74925;

DT 23-APR-2002 (first entry)

DE Synthetic goliath protein sequence #4.

KW Goliath protein; antiangiogenic; vasotropic; gene therapy;
 KW dosage form; angiogenesis; neurogenesis; tumour; vascularisation;
 KW cancer; ischaemia; neuroblastoma; neurodegenerative disease;
 KW unrecovered nerve trauma.

XX Synthetic.

XX WO200193681-A1.

XX 13-DEC-2001.

XX 01-JUN-2001; 2001WO-US18000.

XX 02-JUN-2000; 2000US-0586398.

XX (REGC) UNIV CALIFORNIA.

XX Harland R, Baker JC;

XX WPI; 2002-147637/19.

XX New compositions comprising goliath proteins, useful for modulating
 PT angiogenesis or neurogenesis in mammals e.g. for preventing or
 PT treating undesirable vascularisation of a tumour, ischaemia or
 PT neurodegenerative disease

XX Disclosure; Page 35-36; 45pp; English.

XX The present invention relates to a new pharmaceutical composition that
 CC comprises a goliath polypeptide in dosage form. The goliath polypeptide
 CC has a sequence identity of at least 75% to the protein sequences
 CC (AAU74918-AAU74921) fully defined in the specification. The composition
 CC is useful for modulating angiogenesis or neurogenesis in mammals,
 CC particularly in humans or mice. Specifically, the composition is useful
 CC for the prophylactic and/or therapeutic treatment of excess angiogenesis
 CC e.g. undesirable vascularisation of a tumour or insufficient angiogenesis
 CC e.g. ischaemia, excess neurogenesis e.g. neuroblastoma or insufficient
 CC neurogenesis e.g. neurodegenerative disease or unrecovered nerve trauma.
 CC The present amino acid sequence represents synthetic goliath protein #4
 CC that is one of several artificial goliath proteins (AAU74922-AAU74927)
 CC used in the invention for modulating angiogenesis or neurogenesis.

XX Sequence 428 AA;

Query Match 92.8%; Score 2111; DB 23; Length 428;
 Best Local Similarity 92.8%; Pred. No. 6.9e-219;
 Matches 397; Conservative 11; Mismatches 20; Indels 0; Gaps 0;

QY 1 MGPPPGAGVSCRGCGFRRLLAWCFLLALSPAPGSRGAEAVWTAYLNVSRVPHGTGVR 60

DB 1 MGPPPGIGVYCRGGCGAARLLAWCFLLALSPAPGSRGAEAVWTAYLNVSRVPHGTGVR 60

QY 61 TWELSEGVYQDSPLPVPVAGLVPPDPGALNACNPHNTFTVPTVMGSTVQVSWLALI 120

DB 61 TWELSEGVYQDSPLPVPVAGLVPPDPGALNACNPHNTFTVPTVMGSTVQVSWLALI 120

QY 121 QRGGGCTPADKIHILASERWASGAVIFNPPGTRNEVPMHSHPGAVDIVAIMIGNLKGTKIL 180

DB 121 QRGGGCTPADKIHILASERWASGAVIFNPPGTRNEVPMHSHPGAVDIVAIMIGNLKGTKIL 180

QY 181 QSIQRIQVTWVIEVKGKHGPNWVNHYSIFFVSVPFFITTAATVGYFIFYSARRLRNARAQ 240

DB 181 QSIQRIQVTWVIEVKGKHGPNWVNHYSIFFVSVPFFITTAATVGYFIFYSARRLRNARAQ 240

QY 241 SRKQRLKADAKKATIGRLQLRTLKQDKEIGPDGDSCAVCIELYKPNDLVRLITCNHIFH 300

DB 241 SRKQRLKADAKKATIGRLQLRTLKQDKEIGPDGDSCAVCIELYKPNDLVRLITCNHIFH 300

QY 301 KTCVDPWLLKHKTCPMKCKDILKALGIEVDVEDGVSLSQVPSVNEISNSASSHEEDNRSE 360

DB 301 LTCVDPWLLKHKTCPMKCKDILKALGIEVDVEDGVSLSQVPSVNEISNSASSHEEDNRSE 360

QY 361 TASSGYASVOGTDEPPELHHVQSTNESLQVNVHANSVAVDVIVPHVDNPTFEEDTPOE 420

DB 361 TASSGYASVOGADEPPELHHVQSTNESLQVNVHANSVAVDVIVPHVDNPTFEEDTPOE 420

QY 421 TAVREIKS 428

DB 421 AAVREIKS 428

RESULT 8

AAU74923

ID AAU74923 standard; Protein; 428 AA.

XX

AAU74923;
 23-APR-2002 (first entry)
 Synthetic gliath protein sequence #2.
 Goliath protein; antiangiogenic; vasotropic; gene therapy;
 dosage form; angiogenesis; neurogenesis; tumour; vascularisation;
 cancer; ischaemia; neuroblastoma; neurodegenerative disease;
 unrecovered nerve trauma.
 Synthetic.
 WO200193681-A1.
 13-DEC-2001.
 01-JUN-2001; 2001WO-US18000.
 02-JUN-2000; 2000US-0586398.
 (REGC) UNIV CALIFORNIA.
 Harland R, Baker JC;
 WPI; 2002-147637/19.
 New compositions comprising gliath proteins, useful for modulating
 angiogenesis or neurogenesis in mammals e.g. for preventing or
 treating undesirable vascularisation of a tumour, ischaemia or
 neurodegenerative disease
 Disclosure; Page 33-34; 45pp; English.
 The present invention relates to a new pharmaceutical composition that
 comprises a gliath polypeptide in dosage form. The gliath polypeptide
 has a sequence identity of at least 75% to the protein sequences
 (AAU74918-AAU74921) fully defined in the specification. The composition
 is useful for modulating angiogenesis or neurogenesis in mammals,
 particularly in humans or mice. Specifically, the composition is useful
 for the prophylactic and/or therapeutic treatment of excess angiogenesis
 e.g. undesirable vascularisation of a tumour or insufficient angiogenesis
 e.g. ischaemia, excess neurogenesis e.g. neuroblastoma or insufficient
 neurogenesis e.g. neurodegenerative disease or unrecovered nerve trauma.
 The present amino acid sequence represents synthetic gliath protein #2
 that is one of several artificial gliath proteins (AAU74922-AAU74927)
 used in the invention for modulating angiogenesis or neurogenesis.
 Sequence 428 AA;
 Query Match 92.7%; Score 2110; DB 23; Length 428;
 Best Local Similarity 92.8%; Pred. No. 8.9e-219;
 Matches 397; Conservative 11; Mismatches 20; Indels 0; Gaps 0;
 1 MGPPPGAGVSCRGCGFSRLLAWCFLLALSPQAGSRAEAVMTAYLNVSWRVPHTGVNR 60
 1 MGPPPGIGYVCGGCGAARLLAWCFLLALSPHAFSGRAEAVMTAYLNVAWRVPHTGVNR 60
 61 TWWELSEGVGQDSPLFPVAGLVPPDPGALNACNPHNTFTPTVWGSTVQVSWLALI 120
 61 TWWELSEGVGQDSPLFPVAGLVPPDPGALNACNPHNTFTPTVWGSTVQVSWLALI 120
 121 QRGGCTPADKIHLYERWASGAVTFNPPGTRNEVPMSPHGVADIIVAIMIGNLKGTKIL 180
 121 QRGGCTPADKIHLYERWASGAVTFNPPGTRNEVPMSPHGVADIIVAIMIGNLKGTKIL 180
 181 QSIQRGIQVTWVIEVKKGHPWNHYSIFVSVSFFIITAATVGYFIFYSARRLRNARAQ 240
 181 QSIQRGIQVTWVIEVKKGHPWNHYSIFVSVSFFIITAATVGYFIFYSARRLRNARAQ 240
 241 SRKQRLKADAKAIGRLQLRTLKQDKIEIGPDGSCAVCIELYPNDLVRLTCNHFH 300
 241 SRKQRLKADGKKAIGRLQLRTLKQDKIEIGPDGSCAVCIELYPNDLVRLTCNHFH 300

301 KTCVDPMLLKHTCPMCKCDILKALGIEVDVEDGVSLSQVPSNEISNSASSHEEDNRSE 360
 301 GTCVDPMLLEHRTCPMCKCDILKALGIEVDVEDGVSLSQVPSNEASNTAIPHEEDSRSE 360
 361 TASSGYASVQGTDEBPPLBEHVQSTNESLQVNHANSVAVDVIHVNDPTFEEDETPNOE 420
 361 TASSGYASVQGADEPPLBEHAQSANENLQVNHANSVAVKVPVHDVNDPTFEEDETPNOE 420
 421 TAVREIKS 428
 421 AAVREIKS 428
 RESULT 9
 AAU74926
 ID AAU74926 standard; Protein; 428 AA.
 XX
 AC AAU74926;
 XX
 DT 23-APR-2002 (first entry)
 XX
 DE Synthetic gliath protein sequence #5.
 XX
 KW Goliath protein; antiangiogenic; vasotropic; gene therapy;
 KW dosage form; angiogenesis; neurogenesis; tumour; vascularisation;
 KW cancer; ischaemia; neuroblastoma; neurodegenerative disease;
 KW unrecovered nerve trauma.
 XX
 OS Synthetic.
 XX
 PN WO200193681-A1.
 XX
 PD 13-DEC-2001.
 XX
 PF 01-JUN-2001; 2001WO-US18000.
 XX
 PR 02-JUN-2000; 2000US-0586398.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Harland R, Baker JC;
 XX
 DR WPI; 2002-147637/19.
 XX
 PT New compositions comprising gliath proteins, useful for modulating
 PT angiogenesis or neurogenesis in mammals e.g. for preventing or
 PT treating undesirable vascularisation of a tumour, ischaemia or
 PT neurodegenerative disease
 XX
 PS Disclosure; Page 36-37; 45pp; English.
 XX
 CC The present invention relates to a new pharmaceutical composition that
 CC comprises a gliath polypeptide in dosage form. The gliath polypeptide
 CC has a sequence identity of at least 75% to the protein sequences
 CC (AAU74918-AAU74921) fully defined in the specification. The composition
 CC is useful for modulating angiogenesis or neurogenesis in mammals,
 CC particularly in humans or mice. Specifically, the composition is useful
 CC for the prophylactic and/or therapeutic treatment of excess angiogenesis
 CC e.g. undesirable vascularisation of a tumour or insufficient angiogenesis
 CC e.g. ischaemia, excess neurogenesis e.g. neuroblastoma or insufficient
 CC neurogenesis e.g. neurodegenerative disease or unrecovered nerve trauma.
 CC The present amino acid sequence represents synthetic gliath protein #5
 CC that is one of several artificial gliath proteins (AAU74922-AAU74927)
 CC used in the invention for modulating angiogenesis or neurogenesis.
 XX
 SQ Sequence 428 AA;
 Query Match 92.7%; Score 2108; DB 23; Length 428;
 Best Local Similarity 92.8%; Pred. No. 1.5e-218;
 Matches 397; Conservative 11; Mismatches 20; Indels 0; Gaps 0;
 1 MGPPPGAGVSCRGCGFSRLLAWCFLLALSPQAGSRAEAVMTAYLNVSWRVPHTGVNR 60

Db 1 MGPPPGAGVYCRGGCGAARLLAWCFLLALSPHAPGSRGAEAVWTAYLVNFWRPHTGVNR 60
 QY 61 TWELSEEGVYQDSPLPVPVAGVLPDPGPGALNACNPHNTFTVTVWGSTVQVSWLALI 120
 Db 61 TWELSEEGVYQDSPLPVPVAGVLPDPGPGALNACNPHNTFTVTVWGSTVQVSWLALI 120
 QY 121 QRGGGCTFADKIHLYERWASGAVIFNPFPGTRNEVPMSPHGAVDIVAIMIGNLKGTKIL 180
 Db 121 QRGGGCTFADKIHLYERWASGAVIFNPFPGTRNEVPMSPHGAVDIVAIMIGNLKGTKIL 180
 QY 181 QSTQGIQVMTVIEVCKKHGPMWVHYISIFVSVSPFIITAAATVGYFIFYSARRLRNARQ 240
 Db 181 QSTQGIQVMTVIEVCKKHGPMWVHYISIFVSVSPFIITAAATVGYFIFYSARRLRNARQ 240
 QY 241 SRKQRLKADAKKAIKGLQLRTLKQDKEIGDPGSCAVCIELYKPNLVLRLTCNHIFH 300
 Db 241 SRKQRLKADAKKAIKGLQLRTLKQDKEIGDPGSCAVCIELYKPNLVLRLTCNHIFH 300
 QY 301 KTCVDPWLLKHTCPMCKCDILKALGIEVDVSGSVLQVPVSNISASHEEDNRSE 360
 Db 301 LTCVDPWLLKHTCPMCKCDILKALGIEVDVSGSVLQVPVSNISASHEEDNRSE 360
 QY 361 TASSGVASVQGTDEPPLEHVSQTNESLQLVNHEANSVAVDVI PHVDNPTFEDETPNOE 420
 Db 361 TASSGVASVQGTDEPPLEHVSQTNESLQLVNHEANSVAVDVI PHVDNPTFEDETPNOE 420
 QY 421 TAVREIKS 428
 Db 421 TAVREIKS 428

RESULT 10

AAU74924
 ID AAU74924 standard; Protein; 428 AA.

AC AAU74924;

DT 23-APR-2002 (first entry)

DE Synthetic gliath protein sequence #3.

KW Gliath protein; antiangiogenic; vasotropic; gene therapy;
 KW dosage form; angiogenesis; neurogenesis; tumour; vascularisation;
 KW cancer; ischaemia; neuroblastoma; neurodegenerative disease;
 KW unrecovered nerve trauma.

OS Synthetic.

PN WO200193681-A1.

PD 13-DEC-2001.

PF 01-JUN-2001; 2001WO-US18000.

PR 02-JUN-2000; 2000US-0586398.

XX (REGC) UNIV CALIFORNIA.

XX Harland R, Baker JC;

XX WPI; 2002-147637/19.

XX New compositions comprising gliath proteins, useful for modulating
 PT angiogenesis or neurogenesis in mammals e.g. for preventing or
 PT treating undesirable vascularisation of a tumour, ischaemia or
 PT neurodegenerative disease

XX Disclosure; Page 34-35; 45pp; English.

XX The present invention relates to a new pharmaceutical composition that
 CC comprises a gliath polypeptide in dosage form. The gliath polypeptide
 CC has a sequence identity of at least 75% to the protein sequences

CC (AAU74918-AAU74921) fully defined in the specification. The composition
 is useful for modulating angiogenesis or neurogenesis in mammals,
 CC particularly in humans or mice. Specifically, the composition is useful
 CC for the prophylactic and/or therapeutic treatment of excess angiogenesis
 CC e.g. undesirable vascularisation of a tumour or insufficient angiogenesis
 CC e.g. ischaemia, excess neurogenesis e.g. neuroblastoma or insufficient
 CC neurogenesis e.g. neurodegenerative disease or unrecovered nerve trauma.
 CC The present amino acid sequence represents synthetic gliath protein #3
 CC that is one of several artificial gliath proteins (AAU74922-AAU74927)
 CC used in the invention for modulating angiogenesis or neurogenesis.
 XX

SQ Sequence 428 AA;

Query Match 92.6%; Score 2116; DB 23; Length 428;
 Best Local Similarity 92.8%; Pred. No. 2.4e-218;
 Matches 397; Conservative 10; Mismatches 21; Indels 0; Gaps 0;

QY 1 MGPPPGAGVYCRGGCGFRLAWCFLLALSPHAPGSRGAEAVWTAYLVNFWRPHTGVNR 60
 Db 1 MGPPPGAGVYCRGGCGAARLLAWCFLLALSPHAPGSRGAEAVWTAYLVNFWRPHTGVNR 60
 QY 61 TWELSEEGVYQDSPLPVPVAGVLPDPGPGALNACNPHNTFTVTVWGSTVQVSWLALI 120
 Db 61 TWELSEEGVYQDSPLPVPVAGVLPDPGPGALNACNPHNTFTVTVWGSTVQVSWLALI 120
 QY 121 QRGGGCTFADKIHLYERWASGAVIFNPFPGTRNEVPMSPHGAVDIVAIMIGNLKGTKIL 180
 Db 121 QRGGGCTFADKIHLYERWASGAVIFNPFPGTRNEVPMSPHGAVDIVAIMIGNLKGTKIL 180
 QY 181 QSTQGIQVMTVIEVCKKHGPMWVHYISIFVSVSPFIITAAATVGYFIFYSARRLRNARQ 240
 Db 181 QSTQGIQVMTVIEVCKKHGPMWVHYISIFVSVSPFIITAAATVGYFIFYSARRLRNARQ 240
 QY 241 SRKQRLKADAKKAIKGLQLRTLKQDKEIGDPGSCAVCIELYKPNLVLRLTCNHIFH 300
 Db 241 SRKQRLKADAKKAIKGLQLRTLKQDKEIGDPGSCAVCIELYKPNLVLRLTCNHIFH 300
 QY 301 KTCVDPWLLKHTCPMCKCDILKALGIEVDVSGSVLQVPVSNISASHEEDNRSE 360
 Db 301 LTCVDPWLLKHTCPMCKCDILKALGIEVDVSGSVLQVPVSNISASHEEDNRSE 360
 QY 361 TASSGVASVQGTDEPPLEHVSQTNESLQLVNHEANSVAVDVI PHVDNPTFEDETPNOE 420
 Db 361 TASSGVASVQGTDEPPLEHVSQTNESLQLVNHEANSVAVDVI PHVDNPTFEDETPNOE 420
 QY 421 TAVREIKS 428
 Db 421 TAVREIKS 428

RESULT 11

AAU74924
 ID AAU74924 standard; Protein; 427 AA.

AC AAU74924;

DT 25-MAR-2002 (first entry)

XX Mouse GRAIL.

XX Murine; human; GRAIL; energy; attenuation; tyrosine; phosphorylation;
 KW antigenic stimulation; interleukin-2; gene therapy; polymorphism; IL-2;
 KW autoimmune disease; tumour cell; cancer; transplant rejection.

OS Mus musculus.

XX Key Location/Qualifiers

XX Domain 339..413

XX /label= zinc_RING_finger_domain

XX /note= "Mediates protein/protein interaction"

XX WO200185943-A1.

CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activ/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.

XX SQ Sequence 276 AA;

Query Match 60.0%; Score 1364; DB 22; Length 276;
Best Local Similarity 99.2%; Pred No² 2e-138;
Matches 263; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 164 VDIVAIMGNLKGTKILQSIQRIQVMTVIEVKGKGPWVNHYSIFFVSFFIITAATV 223

Db 12 VDIVAIMGNLKGTKILQSIQRIQVMTVIEVKGKGPWVNHYSIFFVSFFIITAATV 71

QY 224 GYFIFYSARLRNARASQKQKADAKAIGRLQLRTLKQDKEIGPDGDSCAVCIETL 283

Db 72 GYFIFYSARLRNARASQKQKADAKAIGRLQLRTLKQDKEIGPDGDSCAVCIETL 131

QY 284 YKPNDLVRLITCNHIFPKTCVDPWLLKHKTCPMCKDILKALGIEVDVDSVSIQVPVS 343

Db 132 YKPNDLVRLITCNHIFPKTCVDPWLLKHKTCPMCKDILKALGIEVDVDSVSIQVPVS 191

QY 344 NEISNASSHEEDNRSETASSGYASVQGTDEPPLEHVQSTNESIQLVNHANSVAVDVI 403

Db 192 NEISNASSHEEDNRSETASSGYASVQGTDEPPLEHVQSTNESIQLVNHANSVAVDVI 251

QY 404 PHVDNPTFEDETPNQETAVRIKS 428

Db 252 PHVDNPTFEDETPNQETAVRIKS 276

RESULT 13

AAU15897

ID AAU15897 standard; Protein; 225 AA.

XX AC AAU15897;

XX DT 07-NOV-2001 (first entry)

XX DE Human novel secreted protein, Seq ID 850.

XX KW Human; immunosuppressive; antiarthritic; antirheumatic;
KW cytostatic; cardiant; vasotropic; cerebroprotective; nootropic;
KW neuroprotective; antibacterial; virucide; fungicide; ophthalmological;
KW vulnary; secreted protein; rheumatoid arthritis;
KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;
KW cerebrovascular disorder; cerebral ischaemia; angiogenesis;
KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;
KW corneal infection; wound healing; epithelial cell proliferation;
KW skin ageing; food additive; preservative; antiproliferative.

XX OS Homo sapiens.

XX PN WO20015322-A2.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US01341.

XX PR 31-JAN-2000; 2000US-0179065.

XX PR 04-FEB-2000; 2000US-0180628.

XX PR 24-FEB-2000; 2000US-0184664.

XX PR 02-MAR-2000; 2000US-0186350.

XX PR 16-MAR-2000; 2000US-0189874.

XX PR 17-MAR-2000; 2000US-0190076.

XX PR 18-APR-2000; 2000US-0198123.

XX PR 19-MAY-2000; 2000US-0205515.

XX PR 07-JUN-2000; 2000US-0209467.

PR 29-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0217496.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.

PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249219.
 PR 17-NOV-2000; 2000US-0249224.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 (HUMA-) HUMAN GENOME SCI INC.
 Rosen CA, Barash SC, Ruben SM;
 MPI, 2001-488783/53.
 DR N-PSDB; AAS25884.
 New nucleic acid molecules encoding 461 human secreted proteins for
 diagnosing, preventing, treating or ameliorating medical conditions and
 used as food additives or preservatives -
 Claim 11; SEQ ID No 850; 980pp; English.
 The invention relates to isolated nucleic acid molecules and their
 encoded secreted proteins. The nucleic acids and proteins are used to
 prevent, treat or ameliorate a medical condition in e.g. humans, mice,
 rabbits, goats, horses, cats, dogs, chickens or sheep. They
 are also used in diagnosing a pathological condition or susceptibility

CC to a pathological condition. Antibodies to the proteins can also
 CC be used in alleviating symptoms associated with the disorders and in
 CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
 CC immunoassays (ELISA). Disorders which are diagnosed or treated
 CC include autoimmune diseases e.g. rheumatoid arthritis,
 CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
 CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
 CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi
 CC and ocular disorders e.g. corneal infection, and many other
 CC disorders listed in the specification. The polypeptides can also
 CC be used to aid wound healing and epithelial cell proliferation, to
 CC prevent skin aging due to sunburn, to maintain organs before
 CC transplantation, for supporting cell culture of primary tissues, to
 CC regenerate tissues and in chemotaxis. The polypeptides can also be used
 CC as a food additive or preservative to increase or decrease storage
 CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
 CC minerals, cofactors and other nutritional components. The present
 CC sequence represents a novel secreted protein of the invention.

Query Match 49.8%; Score 1134; DB 22; Length 225;
 Best Local Similarity 99.1%; Pred. No. 1.le-113;
 Matches 218; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 209 FVSVSFFIITATGCTFYFYSARRLRNARASRKORQLKADAKKAIQLRLTKQGDK 268
 Db 6 FVSVSFFIITATGCTFYFYSARRLRNARASRKORQLKADAKKAIQLRLTKQGDK 65
 QY 269 EIGPDGDSCAVCIELYKPNLDVRLITCNHIFHKTCDPMLLKHKTCPMCKDILKALGTE 328
 Db 66 EIGPDGDSCAVCIELYKPNLDVRLITCNHIFHKTCDPMLLKHKTCPMCKDILKALGTE 125
 QY 329 VDVEDGSVSLQVPVSNEISNSASSHEEDNRSETASGVASVQGTDEPPLEHVQSTNESL 388
 Db 126 VDVEDGSVSLQVPVSNEISNSASSHEEDNRSETASGVASVQGTDEPPLEHVQSTNESL 185
 QY 389 QLVNHEANSVAVDVIPHVDPNPTFEDETNPQETAVREIKS 428
 Db 186 QLVNHEANSVAVDVIPHVDPNPTFEDETNPQETAVREIKS 225

RESULT 14
 AAU25459
 ID AAU25459 standard; Protein; 294 AA.
 XX AAU25459;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human mddt protein from clone LI:757439.1:2000MAY01.
 XX
 KW Human; molecules for disease detection and treatment; mddt;
 KW Antiarteriosclerotic; hepatotropic; antipsoriatic; cytostatic;
 KW immunosuppressive; antidiabetic; antiasthmatic; neuroprotective;
 KW osteopathic; antiarthritic; cell proliferative disorder;
 KW arteriosclerosis; cirrhosis; psoriasis; cancer; adenocarcinoma;
 KW leukaemia; breast cancer; autoimmune disorder; AIDS;
 KW acquired immunodeficiency syndrome; Addison's disease;
 KW diabetes mellitus; asthma; multiple sclerosis; osteoarthritis.
 XX
 OS Homo sapiens.
 XX
 PN WO200162922-A2.
 XX
 PD 30-AUG-2001.
 XX
 PF 21-FEB-2001; 2001WO-US05896.
 XX
 PR 24-FEB-2000; 2000US-0185213.
 PR 16-MAY-2000; 2000US-0205232.
 PR 17-MAY-2000; 2000US-0205285.
 PR 17-MAY-2000; 2000US-0205286.
 PR 17-MAY-2000; 2000US-0205287.

PR 17-MAY-2000; 2000US-0205323.
 PR 17-MAY-2000; 2000US-0205324.
 XX (INCY-) INCYTE GENOMICS INC.
 XX Panzer SR, Spiro PA, Banville SC, Shah P, Chalup MS, Chang SC;
 PI Chen A, D'Sa SA, Anshey S, Dahl CR, Dam TC, Daniels SE;
 PI Dufour GE, Flores V, Fong WT, Greenwalt LB, Hillman JL, Jones AL;
 PI Liu TF, Roseberry AM, Rosen BH, Russo FD, Stockdreher TK, Daffo A;
 PI Wright RJ, Yap PE, Yu JY, Bradley DL, Bratcher SR, Chen W;
 PI Cohen HJ, Hodgson DM, Lincoln SE, Jackson S;
 XX WPI; 2001-570631/64.
 DR N-PSDB; AAS42511.
 XX
 XX New disease detection and treatment molecule polynucleotides and
 PT polypeptides, useful for diagnosis and treatment of arteriosclerosis,
 PT cirrhosis, psoriasis, cancer, autoimmune disorders, diabetes mellitus,
 PT asthma and multiple sclerosis
 XX
 PS Claim 27; Page 170-171; 183pp; English.
 XX
 CC The invention relates to novel human molecules for disease
 CC detection and treatment (mdt proteins) and the polynucleotides encoding
 CC them. The MDDT polynucleotides and polypeptides are useful for diagnostic
 CC and therapeutic purposes e.g. to diagnose and treat cell proliferative
 CC disorders (e.g. arteriosclerosis, cirrhosis and psoriasis) cancers (e.g.
 CC adenocarcinoma, leukaemia and breast cancer) autoimmune disorders
 CC (e.g. acquired immunodeficiency syndrome (AIDS) and Addison's disease)
 CC diabetes mellitus, asthma, multiple sclerosis, osteoarthritis, and many
 CC more diseases given in the specification. The present sequence
 CC represents an mddt protein of the invention.
 XX
 SQ Sequence 294 AA;
 Query Match 45.2%; Score 1027.5; DB 22; Length 294;
 Best Local Similarity 76.1%; Pred. No. 5.5e-102;
 Matches 201; Conservative 7; Mismatches 25; Indels 31; Gaps 3;
 QY 1 MGPPPGAGVSCRGCGFSLALWCFLLALSPQAPGSRGAEAVWTAYLVNSVRVHTGYNR 60
 DB 43 MGPPPGAGVSCRGCGFSLALWCFLLALSPQAPGSRGAEAVWTAYLVNSVRVHTGYNR 102
 QY 61 TWELSEEGYQDSPLEVAGVLPDPGALNACNPHNTFTVPTVMGSTVQVSWLALI 120
 DB 103 TWELSEEGYQDSPLEVAGVLPDPGALNACNPHNTFTVPTVMGSTVQVSWLALI 162
 QY 121 QRGCGCTFADKTHLAYERVAGAVIFNPPGTENEVTPMSPHGAVDIVAIMIRQSRHKN 177
 DB 163 QRGCGCTFADKTHLAYERVAGAVIFNPPGTENEVTPMSPHGAVDIVAIMIRQSRHKN 222
 QY 178 KILQSGRGIQVTWIEVGKXGHPVWYH-----YSIFFVSVSFFIITAATVG 224
 DB 223 AI-----YSKRHTSDNGHRSREKTPWLGESLNFRRFLCPFLLRRATVG 267
 QY 225 YFIFYSARLRNARQSRQKQK 248
 DB 268 YFIFYSARLRNARQSRQKQK 291
 RESULT 15
 AAU16353
 ID AAU16353 standard; Protein; 165 AA.
 XX
 AC AAU16353;
 XX
 DT 07-NOV-2001 (first entry)
 XX
 DE Human novel secreted protein, Seq ID 1306.
 XX
 KW Human; immunosuppressive; antiarthritic; antirheumatic;
 KW cytostatic; cardiant; vasotropic; cerebroprotective; nootropic;
 KW neuroprotective; antibacterial; virucide; fungicide; ophthalmological;

KW vulnery; secreted protein; rheumatoid arthritis;
 KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;
 KW cerebrovascular disorder; cerebral ischaemia; angio genesis;
 KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;
 KW corneal infection; wound healing; epithelial cell proliferation;
 KW skin ageing; food additive; preservative; antiproliferative.
 OS Homo sapiens.
 XX
 XX WO200155322-A2.
 XX
 XX 02-AUG-2001.
 XX
 XX 17-JAN-2001; 2001WO-US01341.
 XX
 XX 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0218880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225477.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226688.
 PR 23-AUG-2000; 2000US-0227182.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.

PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 29-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.

PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-488783/53.
XX DR N-PSDB; AAS26340.
XX
PT New nucleic acid molecules encoding 461 human secreted proteins for
diagnosing, preventing, treating or ameliorating medical conditions and
used as food additives or preservatives -
XX
PS Claim 11; SEQ ID No 1306; 980pp; English.
XX

The invention relates to isolated nucleic acid molecules and their encoded secreted proteins. The nucleic acids and proteins are used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used in diagnosing a pathological condition or susceptibility to a pathological condition. Antibodies to the proteins can also be used in alleviating symptoms associated with the disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA). Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection, and many other disorders listed in the specification. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors and other nutritional components. The present sequence represents a novel secreted protein of the invention.

Query Match 36.8%; Score 838; DB 22; Length 165;

Best Local Similarity 97.5%; Pred. No. 6.6e-82;

Matches 157; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 268 KEIGPDGDSCAVCIELYKPNLVRILTCNHFHKTCVDPWLLKHKTCPMCKCDILKALGI 327

Db 5 RXIGPDGDSCAVCIELYKPNLVRILTCNHFHKTCVDPWLLKHKTCPMCKCDILKALGI 64

QY 328 EVDVEDGSVLSQVPVNSNISASSSHEEDNRSETASSGVASVQGTDEPPELHVOSTNES 387

DE 65 EVDVEDGSVLSQVPVNSNISASSSHEEDNRSETASSGVASVQGTDEPPELHVOSTNES 124

QY 388 LQLVNHANSVADVIPHVDNPTFEDETPNQETAVERIKS 428

Db 125 LQLVNHANSVADVIPHVDNPTFEDETPNQETAVERIKS 165

Search completed: April 22, 2003, 08:25:54

Job time : 78 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 08:22:11 ; Search time 35 Seconds
(without alignments)
2519.662 Million cell updates/sec

Title: US-09-854-300-8

Perfect score: 2275

Sequence: 1 MGPPPGAGVSCRCGCGFSRL.....PTREEDTPTNQETAVERIKS 428

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2253	99.1	428	Q8TEB7	Q8TEB7 homo sapien
2	2215	97.3	428	Q86RF3	Q86RF3 homo sapien
3	2160	94.9	428	Q9D304	Q9D304 mus musculus
4	2159	94.9	428	Q9JUF8	Q9JUF8 mus musculus
5	2149	94.5	428	Q9DBN3	Q9DBN3 mus musculus
6	1218	53.5	237	Q9CVG1	Q9CVG1 mus musculus
7	930	40.9	376	Q95JW9	Q95JW9 macaca fasc
8	927	40.7	376	Q95K04	Q95K04 macaca fasc
9	910.5	40.0	376	Q8WVZ7	Q8WVZ7 homo sapien
10	829	36.4	316	Q9D3V4	Q9D3V4 mus musculus
11	714	31.4	156	Q9H5E4	Q9H5E4 homo sapien
12	650.5	28.6	426	Q8WU03	Q8WU03 homo sapien
13	595.5	26.2	419	Q8VEM1	Q8VEM1 mus musculus
14	592.5	26.0	419	Q9QZQ6	Q9QZQ6 mus musculus
15	586	25.8	462	Q9ULK6	Q9ULK6 homo sapien
16	468.5	20.6	276	Q9P0J9	Q9P0J9 homo sapien

17	412.5	18.1	461	5	Q8TOM7	Q8TOM7 drosophila
18	356.5	15.7	180	11	Q9CY01	Q9CY01 mus musculus
19	346	15.2	286	5	Q9W0W3	Q9W0W3 drosophila
20	327	14.4	473	5	Q76671	Q76671 caenorhabdi
21	227.5	10.0	310	10	Q9M622	Q9M622 arabidopsis
22	212	9.3	347	11	Q91XF4	Q91XF4 mus musculus
23	211.5	9.3	783	4	Q9NXXD	Q9NXXD homo sapien
24	210	9.2	350	4	Q9H6Y7	Q9H6Y7 homo sapien
25	210	9.2	536	5	Q9V120	Q9V120 drosophila
26	208	9.1	537	10	Q8S7T8	Q8S7T8 oryza sativ
27	205	9.0	424	10	Q9C8W4	Q9C8W4 arabidopsis
28	205	9.0	448	10	Q8VZ14	Q8VZ14 arabidopsis
29	204.5	9.0	431	10	Q9M0N7	Q9M0N7 arabidopsis
30	203	8.9	444	10	Q9M621	Q9M621 arabidopsis
31	203	8.9	456	5	Q9XX98	Q9XX98 caenorhabdi
32	194.5	8.5	324	4	Q9Y3V1	Q9Y3V1 homo sapien
33	191	8.4	891	4	Q9ULT6	Q9ULT6 homo sapien
34	190	8.4	378	10	Q64763	Q64763 arabidopsis
35	189.5	8.3	354	10	Q8S7D2	Q8S7D2 oryza sativ
36	185	8.1	332	10	Q9LQW2	Q9LQW2 arabidopsis
37	184.5	8.1	508	10	Q9C804	Q9C804 arabidopsis
38	184	8.1	404	10	Q9C7T5	Q9C7T5 arabidopsis
39	182.5	8.0	226	11	Q9CYM6	Q9CYM6 mus musculus
40	177	7.8	420	4	Q75866	Q75866 homo sapien
41	177	7.8	420	4	Q8WVF5	Q8WVF5 homo sapien
42	176.5	7.8	362	10	Q9LFF4	Q9LFF4 arabidopsis
43	172	7.6	291	10	Q9SRQ8	Q9SRQ8 arabidopsis
44	172	7.6	322	10	Q9SMQ0	Q9SMQ0 arabidopsis
45	172	7.6	356	10	Q9SS59	Q9SS59 arabidopsis

ALIGNMENTS

RESULT 1

Q8TEB7
ID Q8TEB7 PRELIMINARY; PRT; 428 AA.
AC Q8TEB7
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE CDNA FLJ23684 fis, clone HEP09821.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Oka T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT "NEDO human cDNA sequencing project."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK074264; BAB85033.1; -.
SQ SEQUENCE 428 AA; 46520 MW; 32F9CDB32BF208FA CRC64;

Query Match	99.1%;	Score 2255;	DB 4;	Length 428;
Best Local Similarity	99.3%;	Pred. No. 9.3e-196;		
Matches	425;	Conservative	2;	Mismatches 1;
		Indels	0;	Gaps 0;
QY	1	MGPPPGAGVSCRCGCGFSRL	ALLAWCFLLALSPQAPGSRGAEAVMTAYLVNVS	WRVPHHTGVNR 60
Db	1	MGPPPGAGVSCRCGCGFSRL	ALLAWCFLLALSPQAPGSRGAEAVMTAYLVNVS	WRVPHHTGVNR 60
QY	61	TWELSEGVYQDPLEPVAGVLVPPDPG	ALNACNPHNTFTVPTWGS	TQVSWLAI 120
Db	61	TWELSEGVYQDPLEPVAGVLVPPDPG	ALNACNPHNTFTVPTWGS	TQVSWLAI 120
QY	121	QRGGGCTFADKIHLYERWASGAVL	FNPGTNEVIMPSHPCAVDIVAIMIGNL	KTGL 180
Db	121	QRGGGCTFADKIHLYERWASGAVL	FNPGTNEVIMPSHPCAVDIVAIMIGNL	KTGL 180
QY	181	QSTQRGIQVTWIEVKGKHPWNVHYS	IFFVSVSFFIATAVTGVIFFYSARRLR	NARQA 240

```

Db 181 QSIQRGIQVTVVIEVKKGKGPVWVNHYSIFFVSVSFFIITATVGYFIFYSARLRNARAQ 240
Qy 241 SRKORLKADAKKAIGRLQLRTLKQDKEIGPDGDSCAVCIELYKPNLVRILTCNHIFH 300
Db 241 SRKORLKADAKKAIGRLQLRTLKQDKEIGPDGDSCAVCIELYKPNLVRILTCNHIFH 300
Qy 301 KTCVDPMLLKHKTCPMKCDILKALGIEVDVDSVLSQVPVNSINSASSHEEDNRSE 360
Db 301 KTCVDPMLLKHKTCPMKCDILKALGIEVDVDSVLSQVPVNSINSASSHEEDNRSE 360
Qy 361 TASSGYASVQGTDEPPLEEHVQSTNESLQVNHANSVAVDVIPHVDNPTFEEDTPNQE 420
Db 361 TASSGYASVQGTDEPPLEEHVQSTNESLQVNHANSVAVDVIPHVDNPTFEEDTPNQE 420
Qy 421 TAVREIKS 428
Db 421 TAVREIKS 428

RESULT 2
ID Q96RF3 PRELIMINARY; PRT; 428 AA.
AC Q96RF3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 1300002C13Rik protein (RIKEN CDNA 1300002C13 gene) (Zinc ring
finger-containing protein GRAIL).
GN GRAIL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Soares L.R., Fatman G.;
RT "Grail: a novel ring finger protein upregulated in anergic T cells.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; AF394689; AAK77554.1; -.
DR InterPro; IPR003137; PA.
DR Pfam; PF02225; PA; 1.
DR Zinc-finger.
KW Zinc-finger.
SQ SEQUENCE 428 AA; 46491 MW; E003DBED8968E45A CRC64;

Query Match 97.3%; Score 2213; DB 4; Length 428;
Best Local Similarity 97.7%; Pred. No. 5.9e-192;
Matches 418; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 1 MGPPPGAGVSCRCGCGFSRLLANCFLLALSPQAPGSRGAEAVMTAYLNSWRVPHGTGVR 60
Db 1 MGPPPGAGVSCRCGCGFSRLLANCFLLALSPQAPGSRGAEAVMTAYLNSWRVPHGTGVR 60
Qy 61 TWELSEGVYQDSPLEPVGVLVPPDGPALNACNPHNTFTVPTVWGSTGVSMALI 120
Db 61 TWELSEGVYQDSPLEPVGVLVPPDGPALNACNPHNTFTVPTVWGSTGVSMALI 120
Qy 121 QRGCGCTFADKIHLAVERWAGAVIFNPGTRNEVPMSPHGPADVAIMIGNLKGTKIL 180
Db 121 QRGCGCTFADKIHLAVERWAGAVIFNPGTRNEVPMSPHGPADVAIMIGNLKGTKIL 180
Qy 181 QSIQRGIQVTVVIEVKKGKGPVWVNHYSIFFVSVSFFIITATVGYFIFYSARLRNARAQ 240
Db 181 QSIQRGIQVTVVIEVKKGKGPVWVNHYSIFFVSVSFFIITATVGYFIFYSARLRNARAQ 240
Qy 241 SRKORLKADAKKAIGRLQLRTLKQDKEIGPDGDSCAVCIELYKPNLVRILTCNHIFH 300
Db 241 SRKORLKADAKKAIGRLQLRTLKQDKEIGPDGDSCAVCIELYKPNLVRILTCNHIFH 300
Qy 301 KTCVDPMLLKHKTCPMKCDILKALGIEVDVDSVLSQVPVNSINSASSHEEDNRSE 360

```

```

Db 301 KTCVDPMLLKHKTCPMKCDILKALGIEVDVDSVLSQVPVNSINSASSHEEDNRSE 360
Qy 361 TASSGYASVQGTDEPPLEEHVQSTNESLQVNHANSVAVDVIPHVDNPTFEEDTPNQE 420
Db 361 TASSGYASVQGTDEPPLEEHVQSTNESLQVNHANSVAVDVIPHVDNPTFEEDTPNQE 420
Qy 421 TAVREIKS 428
Db 421 TAVREIKS 428

RESULT 3
ID Q9D304 PRELIMINARY; PRT; 428 AA.
AC Q9D304;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 1300002C13Rik protein (RIKEN CDNA 1300002C13 gene) (Zinc ring
finger-containing protein GRAIL).
GN GRAIL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schrim L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:695-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=BREAST TUMOR;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=DGA/2;
RA Ford G.S., Bloom D., Paragas V., Anandasabapathy N., Skrenta H.,
RA Ermann J., Cron R.Q., Lewis D.B., Fatman G.;
RT "GRAIL: A novel gene expressed in anergic T cells that inhibits
NFAT/AP-1 activation and IL-2 transcription.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; AK018582; BAB31291.1; -.
DR EMBL; BC010477; AAH10477.1; -.
DR EMBL; AF426411; AAL34514.1; -.
DR MGI; MGI:1914139; 1300002C13Rik.
DR InterPro; IPR003137; PA.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF02225; PA; 1.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
KW Zinc-finger.

```

SQ SEQUENCE 428 AA; 46276 MW; 480DCP46CT5E238F CRC64;

Query Match 94.9%; Score 2160; DB 11; Length 428;
 Best Local Similarity 94.6%; Pred. No. 3.8e-187;
 Matches 405; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

QY 1 MGPPPGAGVSCRGCGFSRLLAFCFLALSPQAPGSRGAEAVMTAYLVNSWRVPHTGVNR 60
 DB 1 MGPPPGIGVYCRGGCGAARLLAFCFLALSPHAPGSRGAEAVMTAYLVNSWRVPHTGVNR 60

QY 61 TWELSEEGVYGQDSPLEPVAGVLVPPDPGALNACNPHNTFTVPTVMGSTVQVSWLALI 120
 DB 61 TWELSEEGVYGQDSPLEPVAGVLVPPDPGALNACNPHNTFTVPTVMGSTVQVSWLALI 120

QY 121 ORGGGCTFADKTHLAYERWASGAVIFNFPGTRNEVTPMSHPGADVIVAIMGNLKGTKIL 180
 DB 121 ORGGGCTFADKTHLAYERWASGAVIFNFPGTRNEVTPMSHPGADVIVAIMGNLKGTKIL 180

QY 181 QSIQGIQVTWIEVGKKGPMWVHYISFFVSFFIITAAATGVYFYSARRLRNRAQ 240
 DB 181 QSIQGIQVTWIEVGKKGPMWVHYISFFVSFFIITAAATGVYFYSARRLRNRAQ 240

QY 241 SRKQRLKADAKAIGRLQLRTLKQDKEIGPDGSCAVCIELYKPNLVLRLTCNHIFH 300
 DB 241 SRKQRLKADAKAIGRLQLRTLKQDKEIGPDGSCAVCIELYKPNLVLRLTCNHIFH 300

QY 301 KTCVDPMLLKHKTCPMCKDILKALGIEVDVDSVSLQVPVSNESASHEEDNRSE 360
 DB 301 KTCVDPMLLKHKTCPMCKDILKALGIEVDVDSVSLQVPVSNESASHEEDNRSE 360

QY 361 TASSGVASVQGTDEPPLEHVSQTNESLQLVNHEANSVAVDVIPHVDNPTFEDETPNOE 420
 DB 361 TASSGVASVQGTDEPPLEHVSQTNESLQLVNHEANSVAVDVIPHVDNPTFEDETPNOE 420

QY 421 TAVREIKS 428
 DB 421 AAVREIKS 428

RESULT 4

Q9JUF8 ID Q9JUF8 PRELIMINARY; PRT; 428 AA.

AC Q9JUF8
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE Brain cDNA, clone MNCB-3816, similar to AF171875 gl-related zinc finger protein (Mus musculus).
 GN 1300002C13RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RA Osada N., Kusuda J., Tanuma R., Ito A., Hirata M., Sugano S., Hashimoto K.
 RT "Isolation of full-length cDNA clones from mouse brain cDNA library made by oligo-capping method."
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 CC -|- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
 DR EMBL; AB041548; BAA95033.1; -
 DR MGD; MGI:1914139; 1300002C13RIK.
 DR InterPro; IPR003137; PA.
 DR Pfam; PF02225; PA; 1.
 DR SMART; SM00097; zf-C3HC4; 1.
 DR SMART; SM00184; RING; 1.
 KW Zinc-finger.
 SQ SEQUENCE 428 AA; 46262 MW; E13AGDCAD67DECBDC CRC64;

Query Match 94.9%; Score 2159; DB 11; Length 428;

Best Local Similarity 94.4%; Pred. No. 4.6e-187;
 Matches 404; Conservative 11; Mismatches 13; Indels 0; Gaps 0;

QY 1 MGPPPGAGVSCRGCGFSRLLAFCFLALSPQAPGSRGAEAVMTAYLVNSWRVPHTGVNR 60
 DB 1 MGPPPGIGVYCRGGCGAARLLAFCFLALSPHAPGSRGAEAVMTAYLVNSWRVPHTGVNR 60

QY 61 TWELSEEGVYGQDSPLEPVAGVLVPPDPGALNACNPHNTFTVPTVMGSTVQVSWLALI 120
 DB 61 TWELSEEGVYGQDSPLEPVAGVLVPPDPGALNACNPHNTFTVPTVMGSTVQVSWLALI 120

QY 121 ORGGGCTFADKTHLAYERWASGAVIFNFPGTRNEVTPMSHPGADVIVAIMGNLKGTKIL 180
 DB 121 ORGGGCTFADKTHLAYERWASGAVIFNFPGTRNEVTPMSHPGADVIVAIMGNLKGTKIL 180

QY 181 QSIQGIQVTWIEVGKKGPMWVHYISFFVSFFIITAAATGVYFYSARRLRNRAQ 240
 DB 181 QSIQGIQVTWIEVGKKGPMWVHYISFFVSFFIITAAATGVYFYSARRLRNRAQ 240

QY 241 SRKQRLKADAKAIGRLQLRTLKQDKEIGPDGSCAVCIELYKPNLVLRLTCNHIFH 300
 DB 241 SRKQRLKADAKAIGRLQLRTLKQDKEIGPDGSCAVCIELYKPNLVLRLTCNHIFH 300

QY 301 KTCVDPMLLKHKTCPMCKDILKALGIEVDVDSVSLQVPVSNESASHEEDNRSE 360
 DB 301 KTCVDPMLLKHKTCPMCKDILKALGIEVDVDSVSLQVPVSNESASHEEDNRSE 360

QY 361 TASSGVASVQGTDEPPLEHVSQTNESLQLVNHEANSVAVDVIPHVDNPTFEDETPNOE 420
 DB 361 TASSGVASVQGTDEPPLEHVSQTNESLQLVNHEANSVAVDVIPHVDNPTFEDETPNOE 420

QY 421 TAVREIKS 428
 DB 421 AAVREIKS 428

RESULT 5

Q9DBN3 ID Q9DBN3 PRELIMINARY; PRT; 428 AA.

AC Q9DBN3
 DT 01-JUN-2001 (TRENBLrel. 17, Created)
 DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE 1300002C13RIK protein.
 GN 1300002C13RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=LIVER;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana S., Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kuchiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schriml L.M., Staubli P., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Marzari J., Mombarts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L., Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohseki S., Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).

CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
 DR EMBL; AK004847; BAB23613.1; -;
 DR MGD; MGI:1914139; 1300002C13Rik.
 DR InterPro; IPR003137; PA.
 DR InterPro; IPR001841; Znf_ring.
 DR Pfam; PF02225; PA; 1.
 DR Pfam; PF00097; zf-C3HC4; 1.
 DR SMART; SM00184; RING; 1.
 KW Zinc-finger.
 SQ SEQUENCE 428 AA; 46292 MW; EBA2B505218288B8 CRC64;

Query Match 94.5%; Score 2149; DB 11; Length 428;
 Best Local Similarity 94.2%; Pred. No. 3.7e-186;
 Matches 403; Conservative 10; Mismatches 15; Indels 0; Gaps 0;
 QY 1 MGPPPGAGVSCRGCGFSRLAWCFLLALSPQAPGSRGAETAYLVNSWRVPHGTGVR 60
 DB 1 MGPPPGIGVYCRGCGAARLLAWCFLLALSPHAPGSRGAETAYLVNSWRVPHGTGVR 60
 QY 61 TWELSEEGVYGDSPLPVPVPPDGGALNACNPHNTFTVPTVMGSTVQVSWLALI 120
 DB 61 TWELSEEGVYGDSPLPVPVPPDGGALNACNPHNTFTVPTVMGSTVQVSWLALI 120
 QY 121 QRGCGCTFADKHLAYERWASGAVIFNFGTRNEVPMSPHAGAVIVAIMIGNLKGTKIL 180
 DB 121 QRGCGCTFADNHLASERGASGAVIFNFGTRNEVPMSPHAGAVIVAIMIGNLKGTKIL 180
 QY 181 QSTQGIQVTWIEVGKGGPWNHYISIFVSVSPFIITAAATGVGFYFISARLRNARQA 240
 DB 181 QSTQGIQVTWIEVGKGGPWNHYISIFVSVSPFIITAAATGVGFYFISARLRNARQA 240
 QY 241 SRKROLKADAKKAIGRLQLRTLKQDGEIGPDGSCAVCIELYPNDLVRILTCNHIFH 300
 DB 241 SRKROLKADAKKAIGRLQLRTLKQDGEIGPDGSCAVCIELYPNDLVRILTCNHIFH 300
 QY 301 KTCVDPLWKHKTCPCMKCDILKALGIEVDVDEGVSLOVPVNSISASHEEDNSE 360
 DB 301 KTCVDPLWLEHRTCPMKCDILKALGIEVDVDEGVSLOVPVNSASNTASPEEDNSE 360
 QY 361 TASSGYASVQGTDEPPLEEHVOSTNESLQLVNHEANSVAVDVIPIVDNPTFDEETPQOE 420
 DB 361 TASSGYASVQGTDEPPLEEHVOSTNESLQLVNHEANSVAVDVIPIVDNPTFDEETPQOE 420
 QY 421 TAVREIKS 428
 DB 421 AAVREIKS 428

RESULT 6
 Q9CVG1 ID Q9CVG1 PRELIMINARY; PRT; 237 AA.
 AC Q9CVG1
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DE 1300002C13Rik protein (Fragment).
 GN 1300002C13Rik.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=SMALL INTESTINE;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadori K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Glass C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai J., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake K., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bulc C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazarrelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayaishizaki Y.;
 RT *Functional annotation of a full-length mouse cDNA collection.*;
 RL Nature 409:685-690(2001).
 DR EMBL; AK008312; BAB25595.1; -;
 DR MGD; MGI:1914139; 1300002C13Rik.
 DR InterPro; IPR003137; PA.
 DR Pfam; PF02225; PA; 1.
 FT NON_TER 237
 SQ SEQUENCE 237 AA; 25344 MW; 319E6AA548F7379B CRC64;
 Query Match 53.5%; Score 1218; DB 11; Length 237;
 Best Local Similarity 96.2%; Pred. No. 3.8e-102;
 Matches 228; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
 QY 1 MGPPPGAGVSCRGCGFSRLAWCFLLALSPQAPGSRGAETAYLVNSWRVPHGTGVR 60
 DB 1 MGPPPGIGVYCRGCGAARLLAWCFLLALSPHAPGSRGAETAYLVNSWRVPHGTGVR 60
 QY 61 TWELSEEGVYGDSPLPVPVPPDGGALNACNPHNTFTVPTVMGSTVQVSWLALI 120
 DB 61 TWELSEEGVYGDSPLPVPVPPDGGALNACNPHNTFTVPTVMGSTVQVSWLALI 120
 QY 121 QRGCGCTFADKHLAYERWASGAVIFNFGTRNEVPMSPHAGAVIVAIMIGNLKGTKIL 180
 DB 121 QRGCGCTFADKHLASERGASGAVIFNFGTRNEVPMSPHAGAVIVAIMIGNLKGTKIL 180
 QY 181 QSTQGIQVTWIEVGKGGPWNHYISIFVSVSPFIITAAATGVGFYFISARLRNARQA 237
 DB 181 QSTQGIQVTWIEVGKGGPWNHYISIFVSVSPFIITAAATGVGFYFISARLRNARQA 237

RESULT 7
 Q95JW9 ID Q95JW9 PRELIMINARY; PRT; 376 AA.
 AC Q95JW9
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Hypothetical 42.2 kDa protein.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
 OC Cercopitheidae; Macaca.
 OX NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=TESTIS;
 RA Hashimoto K., Osada N., Hida M., Kusuda J., Tanuma R., Hirai M.,
 RA Terao K., Sugano S.;
 RT "Isolation of novel full-length cDNA clones from macaque testis cDNA
 RL libraries."
 RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
 DR EMBL; AB070058; BAB63003.1; -;
 DR InterPro; IPR001309; ICE_P20.
 DR InterPro; IPR003137; PA.
 DR InterPro; IPR001841; Znf_ring.
 DR Pfam; PF02225; PA; 1.
 DR Pfam; PF00097; zf-C3HC4; 1.
 DR PROSITE; PS01121; CASPASE_HIS; UNKNOWN_1.
 KW Hypothetical protein; Zinc-finger.
 SQ SEQUENCE 376 AA; 42248 MW; B349791C19956841 CRC64;
 Query Match 40.9%; Score 930; DB 6; Length 376;

Best Local Similarity 48.6%; Pred. No. 9.4e-76;
Matches 191; Conservative 65; Mismatches 109; Indels 28; Gaps 10;

```

QY 12 RGCGRSLLAWCFLLALSPQAGSRGAEAVTAYLVNSWRVPHGTGNRTWELSEEGVY 71
Db 10 RNTAFSLWIMFGVWLWVSQNC---CRASVVMWTAYNLSF---HVG-NHVLSEGETGVF 62

QY 72 GDSPLPVPAGVLPVPPDGFALNACNPHNTFTVPTVMGSTVQVSWLAIQIRGGGCTPADK 131
Db 63 GRSTLKR VAGVLPVPEGK-IQNACNPNTIFS-----RSKYSETWALIER-GGCTFTQK 115

QY 132 IHLAYERWASGAVINFPPGTRNEVPMSPHGAVDIVAIMIGNLKTGLQSIQIRGIQVTM 191
Db 116 IKVAAEKASGVIIYNNPPTGNQVFPMPHQAFEDVVMVIMGNLKGTEIFHLIKKGLVITA 175

QY 192 VIEVGKKGHPWNHYSIFFVSVSFFIITAAATVGYEIFYSARLRNARAKRQKOLKADA 251
Db 176 MVEVGRKHIIWMNHY-----LVSFVITATLAYFIYHRLCLARIQNRWQRLTDL 230

QY 252 KKAIGRLQRLTKQSGKEIGPDGSCAVCIELYKPNLDVRLTNCNHFHKTCTVDPLWLKH 311
Db 231 ONAFGQLQRLVVKEGDEEINPNNGSCVICFELYKPNLDIVRLTCKHFFKNCIDPWILSH 290

QY 312 KTCPCMKCDILKALGIEVDVDSVSLQVPSNEISNGASHEEDNRSETASSGYASVOG 371
Db 291 GTCPCICKCDILKVLGQVDVNGTBPLOVMSSELCETLSPSEETNNE-----VSPAG 344

QY 372 TDEPPLERHVSQTNESLQVLNHEANSVAVDVIP 404
Db 345 TSDKVI--HVEE-NPTSQNDSQPHSVVEDVHP 374

```

RESULT 8

```

Q95K04
ID Q95K04 PRELIMINARY; PRT; 376 AA.
AC Q95K04;
DT 01-DEC-2001 (TREMELrel. 19, Created)
DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMELrel. 20, Last annotation update)
DE Hypothetical 42.2 kDa protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA Hashimoto K., Osada N., Hida M., Kusuda J., Tanuma R., Hirai M.,
RA Terao K., Sugano S.
RT "Isolation of novel full-length cDNA clones from macaque testis cDNA
RT libraries."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; AB070023; BAB62968.1; -.
DR InterPro; IPR001309; ICE_p20.
DR InterPro; IPR001337; PA.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF02225; PA; 1.
DR Pfam; PF00097; zf-C3HC4; 1.
DR PROSITE; PS01121; CASPASE_HIS; UNKNOWN_1.
KW Hypothetical protein; ZINC-finger.
SQ SEQUENCE 376 AA; 42238 MW; D44585F36F4E0C6E CRC64;

```

Query Match 40.7%; Score 927; DB 6; Length 376;
Best Local Similarity 48.6%; Pred. No. 1.8e-75;
Matches 191; Conservative 65; Mismatches 109; Indels 28; Gaps 10;

```

QY 12 RGCGRSLLAWCFLLALSPQAGSRGAEAVTAYLVNSWRVPHGTGNRTWELSEEGVY 71
Db 10 RNTAFSLWIMFGVWLWVSQNC---CRASVVMWTAYNLSF---HVG-NHVLSEGETGVF 62

QY 72 GDSPLPVPAGVLPVPPDGFALNACNPHNTFTVPTVMGSTVQVSWLAIQIRGGGCTPADK 131

```

```

Db 63 GRSTLKR VAGVLPVPEGK-IQNACNPNTIFS-----RSKYSETWALIER-GGCTFTQK 115

QY 132 IHLAYERWASGAVINFPPGTRNEVPMSPHGAVDIVAIMIGNLKTGLQSIQIRGIQVTM 191
Db 116 IKVAAEKASGVIIYNNPPTGNQVFPMPHQAFEDVVMVIMGNLKGTEIFHLIKKGLVITA 175

QY 192 VIEVGKKGHPWNHYSIFFVSVSFFIITAAATVGYEIFYSARLRNARAKRQKOLKADA 251
Db 176 MVEVGRKHIIWMNHY-----LVSFVITATLAYFIYHRLCLARIQNRWQRLTDL 230

QY 252 KKAIGRLQRLTKQSGKEIGPDGSCAVCIELYKPNLDVRLTNCNHFHKTCTVDPLWLKH 311
Db 231 ONAFGQLQRLVVKEGDEEINPNNGSCVICFELYKPNLDIVRLTCKHFFKNCIDPWILSH 290

QY 312 KTCPCMKCDILKALGIEVDVDSVSLQVPSNEISNGASHEEDNRSETASSGYASVOG 371
Db 291 GTCPCICKCDILKVLGQVDVNGTBPLOVMSSELCETLSPSEETNNE-----VSPAG 344

QY 372 TDEPPLERHVSQTNESLQVLNHEANSVAVDVIP 404
Db 345 TSDKVI--HVEE-NPTSQNDSQPHSVVEDVHP 374

```

RESULT 9

```

Q8WVZ7
ID Q8WVZ7 PRELIMINARY; PRT; 376 AA.
AC Q8WVZ7;
DT 01-MAR-2002 (TREMELrel. 20, Created)
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)
DE Hypothetical 42.3 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA Strausberg R.
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC022038; AAH22038.1; -.
DR InterPro; IPR003137; PA.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF02225; PA; 1.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00069; ZF_RING_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 376 AA; 42294 MW; 7C0446D1FB0981F1 CRC64;

```

Query Match 40.0%; Score 910.5; DB 4; Length 376;
Best Local Similarity 50.0%; Pred. No. 5.5e-74;
Matches 184; Conservative 58; Mismatches 97; Indels 29; Gaps 9;

```

QY 39 AEAVTAYLVNSWRVPHGTGNRTWELSEEGVQDSPLPVPAGVLPVPPDGFALNACNP 98
Db 34 ASVVMWYMNISF---HVG-NHVLSEGETGVFGRSSTLKR VAGVLPVPEGK-IQNACNP 88

QY 99 HTNFTPTVMGSTVQVSWLAIQIRGGGCTFADKIHLYERWASGAVINFPPGTRNEVPM 158
Db 89 NTIFS-----RSKYSETWALIER-GGCTFTOKIRVATEKGASGVIIYNNPVTGNQVFP 142

QY 159 SHPGAVDIVAIMIGNLKTGLQSIQIRGIQVTMVEVGRKHIIWMNHY-----LVSFVIV 218
Db 143 FHQAFEDVVMVIMGNLKGTEIFHLIKKGLVITAVVEVGRKHIIWMNHY-----LVSFVIV 197

QY 219 TAATVGYEIFYSARLRNARAKRQKOLKADA KKAIGRLQRLTKQSGKEIGPDGSDCA 278
Db 198 TTATLAYFIYHRLCLARIQNRWQRLTDLQNTFGQLQRLVVKEGDEEINPNNGSCV 257

QY 279 VCIELYKPNLDVRLTNCNHFHKTCTVDPLWLKHKTCPMKCDILKALGIEVDVDSVSL 338

```


Db 119 ER -GGCTFTTKHINLAEBKANGVVIIVNYPGTGNKVPMPMHSQGTENIVAVMIGNLKGWELL 171

Qy 181 QSIORGQIVTMVIEVGKGGHPWNHYHSIFPVSVFFIITAAVTGVYFYSARLRNARAQ 240

Db 178 HLIQOQGVYVITIEVGRMHPMLSHY-----VMSLFTFLAAVTVLYFLYCAWRPRVSNSS 232

Qy 241 SRKOROLKADAKKAIGRLQRLTLKQDKEIGPDGDSCAVCIELYKPNDLVRLTCNHIFH 300

Db 233 TRKQRQLKADVKKAIQGLQURVLQDSEKDELDPNEDSCVCFDMYKAQDVIRILTCKHFFH 392

Qy 301 KTCVDPWLLKHKTCPCMKCDILK 323

Db 293 KTCIDPWLIAHRTCPMKCDILK 315

RESULT 11

Q9H5E4

ID Q9H5E4 PRELIMINARY; PRT; 156 AA.

AC Q9H5E4;

DT 01-MAR-2001 (TrEMBLrel. 16, Created)

DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE-CDNA: FLJ23516 fis, clone LMG04848.

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

XP [1]

RP SEQUENCE FROM N.A.

RA TISSUE=LUNG;

RA Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,

RA Nakajima Y., Mizuno T., Morinaga M., Tanigami A., Fujiwara T., Ono T.,

RA Yamada K., Fujii Y., Ozaki K., Hikao M., Ohmori Y., Ota T., Suzuki Y.,

RA Obayashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y.,

RA Iweda T., Sugano S.;

RT "NEDO human cDNA sequencing project.";

RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.

DR EMBL; AK027169; BAB15682.1; -

DR InterPro; IPR001841; Znf_ring.

DR Pfam; PF00097; zf-C3HC4_1.

DR SMART; SM00184; RING; 1.

KW Zinc-finger.

SQ SEQUENCE 156 AA; 17263 MW; 526F427554DD4FB9 CRC64;

Query Match 31.4%; Score 714; DB 4; Length 156;

Best Local Similarity 89.7%; Pred. No. le-56;

Matches 139; Conservative 5; Mismatches 3; Indels 8; Gaps

Qy 277 CAV-CIELYKPNDLV--RLTCHNHFHKTCDPWLKHKHTCPMKCDILKALGIEVDVD 333

Db 7 CALNCI-----NQMIWYSILTCNHIFHKTCDPWLLEHRTCPMKCDILKALGIEVDVD 61

Qy 334 GSVSLQVPVNEISNSASSHEEDNRSETASSGVSQGTDEPPELHEHVOSTNESLQLVNH 393

Db 62 GSASLQVPVNEISNSASSHEEDNRSETASSGVSQGTDEPPELHEHVOSTNESLQLVNH 121

Qy 394 EANSVAVDVIVPHVDNPTFEEDTPNQETAVREIKS 428

Db 122 EANSVAVDVIVPHVDNPTFEEDTPNQETAVREIKS 156

RESULT 12

Q8WU03

ID Q8WU03 PRELIMINARY; PRT; 426 AA.

AC Q8WU03;

DT 01-MAR-2002 (TrEMBLrel. 20, Created)

DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE Hypothetical 45.7 kDa protein (Fragment).

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Best Local Similarity 42.1%; Pred. No. 4.2e-45;
Matches 133; Conservative 50; Mismatches 120; Indels 13; Gaps 7;

QY 16 GFSRLIACFLALSPQAGSABAVMTAVLNVSRRVPHGTGVNRTWELSEEGVYGDS 75
DB 8 GPALAAALALTCSLWPRADNASOEYITALINVTVOEPGRGTPLTF--RIDRGYGLDS 65

QY 76 PLEPVAGVLVPP---DGPALNACNPHNTFTVPTVMGSTVQVSMALIQSGGCTFADKI 132
DB 66 PKAERGVQVLAFLPTIHGVADHLGCDPQTRFFVP-----PNIK-QWIALIQR-GNCTFKEKI 119

QY 133 HLAERWASGAVIENFPPTRENEVPMSPHGAVIDIVAIMIGNLKGTKILQSIQGIQVMTV 192
DB 120 SRAAFHNAVAVIYN-NKSEBPVMTHTPGTGDIIVAVMITELRGDILSYLEKISVQMT 178

QY 193 IEVKGKHPW-VNHYSIFFVSVPFIITAATVGYFIFYSARRLRNARAQSRKQRLKADA 251
DB 179 IAVGTRMPPKPNFSGSLVFSISFVLMIISSAWLIFYFIQIRVYNARDNRQRLGDA 238

QY 252 KKAIGRLQRLTKQDKEIGPDGSCAVCIELYKPNDLVRILTCNHIFHKTCTVDPWLLKH 311
DB 239 KKAISKLTRTVKKGKEDDPDFDHCAVCIESYKQNDVVRVLPCKHVFHKSCVDPWLSEH 298

QY 312 KTCPMCKCDILKALGI 327
DB 299 CTCPMCKNLILRALGI 314

RESULT 15
Q9ULK6 PRELIMINARY; PRT; 462 AA.

AC Q9ULK6; MEDLINE=20039619; PubMed=10574462;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE KIAA1214 protein (Fragment).
GN KIAA1214.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Nagase T., Ishikawa K., Kikuno R., Hirose M., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XV.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro."
RL DNA Res. 6:337-345(1999).
CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL: AB033040; BA86528.1; -.
DR InterPro: IPR003137; PA.
DR InterPro: IPR001841; Znf_ring.
DR Pfam: PF02225; PA; 1.
DR Pfam: PF00097; zf-C3HC4; 1.
DR SMART: SM00184; RING; 1.
KW Zinc-finger.
FT NON_TER
SQ SEQUENCE 462 AA; 49883 MW; D1279DE127C9772D CRC64;

Query Match 25.8%; Score 586; DB 4; Length 462;
Best Local Similarity 35.6%; Pred. No. 1.9e-44;
Matches 156; Conservative 46; Mismatches 124; Indels 112; Gaps 11;

QY 6 GAGVSCRCGGGFSR-----LLAWCF--LLALSPOAGSGRAE 40
DB 47 GGGTGRREGAGAGRPRIKMAISLQACCSLALSTWLLSFCFVHLLCLDFTVAEK---E 103

QY 41 AVMTAYLNVSRRVP-----HTGVNRTWELSEEGVYGDSPLFPVAGVL 84
DB 104 EWYTAFTVITYAEPAPDPGAGAGGGAGELHT-----EKTECGRYGHSFKQDARVE 156

Search completed: April 22, 2003, 08:26:58
Job time : 38 secs